Star formation in semi-analytic galaxy formation models with multiphase gas

Somerville, Rachel S.; Popping, Gergö; Trager, Scott C.

Published in:
Monthly Notices of the Royal Astronomical Society

DOI:
10.1093/mnras/stv1877

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY
MNRAS **453,** 4337–4367 (2015) doi:10.1093/mnras/stv1877

Star formation in semi-analytic galaxy formation models with multiphase gas

Rachel S. Somerville,[1★] Gergö Popping[2,3] and Scott C. Trager[2]

[1]*Department of Physics and Astronomy, Rutgers University, 136 Frelinghuysen Road, Piscataway, NJ 08854, USA*
[2]*Kapteyn Astronomical Institute, University of Groningen, Postbus 800, NL-9700 AV Groningen, the Netherlands*
[3]*European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-85748 Garching, Germany*

Accepted 2015 August 12. Received 2015 August 10; in original form 2015 February 18

ABSTRACT

We implement physically motivated recipes for partitioning cold gas into different phases (atomic, molecular, and ionized) in galaxies within semi-analytic models of galaxy formation based on cosmological merger trees. We then model the conversion of molecular gas into stars using empirical recipes motivated by recent observations. We explore the impact of these new recipes on the evolution of fundamental galaxy properties such as stellar mass, star formation rate (SFR), and gas and stellar phase metallicity. We present predictions for stellar mass functions, stellar mass versus SFR relations, and cold gas phase and stellar mass–metallicity relations for our fiducial models, from redshift $z \sim 6$ to the present day. In addition we present predictions for the global SFR, mass assembly history, and cosmic enrichment history. We find that the predicted stellar properties of galaxies (stellar mass, SFR, metallicity) are remarkably insensitive to the details of the recipes used for partitioning gas into H I and H_2. We see significant sensitivity to the recipes for H_2 formation only in very low mass haloes ($M_{\rm h} \lesssim 10^{10.5}\,{\rm M}_{\odot}$), which host galaxies with stellar masses $m_* \lesssim 10^8\,{\rm M}_{\odot}$. The properties of low-mass galaxies are also quite insensitive to the details of the recipe used for converting H_2 into stars, while the formation epoch of massive galaxies does depend on this significantly. We argue that this behaviour can be interpreted within the framework of a simple equilibrium model for galaxy evolution, in which the conversion of cold gas into stars is balanced on average by inflows and outflows.

Key words: galaxies: evolution – galaxies: formation – galaxies: high-redshift – galaxies: star formation.

1 INTRODUCTION

While the ΛCDM (cold dark matter plus cosmological constant Λ) model (Blumenthal et al. 1984) now provides us with a well-motivated framework for predicting the abundances and properties of dark matter haloes and the large scale structures in which they are embedded, all galactic or larger scale simulations must rely on ‘sub-grid’ recipes in order to treat processes such as star formation (SF) and stellar feedback. Cosmological simulations are unable to directly resolve individual stars or, usually, even Giant Molecular Clouds (GMCs). In order to model the conversion of cold gas into stars, up until recently, both numerical and semi-analytic cosmological simulations typically utilized a very simple empirical sub-grid recipe based on observations most famously by Schmidt (1959, 1963) and Kennicutt (1989, 1998) (often referred to as the ‘Kennicutt–Schmidt’ (KS) relation). These observations showed that the surface density of SF $\Sigma_{\rm SFR}$ was proportional to the surface density of cold gas to a power $N_{\rm KS}$. Observations also showed that the efficiency of SF dropped rapidly below a critical gas surface density (Martin & Kennicutt 2001). There has been debate about whether this critical surface density is best described in terms of a Toomre stability criterion (Toomre 1964) or a constant critical density, and indeed about the physical origin of this critical density (Schaye 2004; Leroy et al. 2008).

From the pioneering work of Katz (1992) up until recently, cosmological simulations of galaxy formation, both numerical and semi-analytic, have implemented an SF recipe in which ‘cold’ gas (typically with $T \lesssim 10^4$ K) with volume density $\rho_{\rm gas}$ is assumed to form stars at a rate per unit volume:

$$\dot{\rho}_* = \epsilon_* \rho_{\rm gas}^N, \qquad (1)$$

with $N \simeq 1.5$ and ϵ_* usually treated as a free parameter, tuned to match the observed Kennicutt relation. A common variant assumes

★ E-mail: somerville@physics.rutgers.edu

$\dot{\rho}_* \propto \rho_{gas}/t_{ff}$, which is approximately equivalent because the local free-fall time $t_{ff} \propto \rho^{-0.5}$. Motivated by the observational evidence described above, many modellers incorporated either a critical surface density or volume density into their SF recipe, which proved to be important in order to reproduce the observed high gas fractions in low-mass galaxies.

Beginning about a decade ago, our understanding of how SF on ∼100 pc–kpc scales depends on local conditions started to undergo a revolution. Wong & Blitz (2002) showed that the correlation between Σ_{SFR} and the surface density of *molecular* gas Σ_{H_2} was stronger than that between Σ_{SFR} and the *total* gas density Σ_{gas} in molecule rich galaxies. In the past five years, this field has advanced rapidly with the availability of galaxy-wide, high-resolution maps of the SF and multiphase (H I and H_2) gas in reasonably large samples of nearby galaxies, e.g. from the THINGS (The H I nearby galaxy survey; Walter et al. 2008) combined with BIMA SONG (BIMA survey of Nearby Galaxies; Helfer et al. 2003) and HERACLES (HERA CO-Line Extragalactic Survey; Leroy et al. 2009). Based on these observations, it has been shown (Bigiel et al. 2008, 2011; Leroy et al. 2008; Schruba et al. 2011) that, when averaged over scales of ∼700 pc, the SF density is tightly correlated with the *molecular* gas density to a nearly linear power, and that there is almost *no* correlation between Σ_{SFR} and the density of atomic gas, so that the correlation between Σ_{SFR} and Σ_{gas} (the traditional KS relation) breaks down badly in the H I-dominated parts of galaxies (typically in galaxy outskirts). These results highlight the importance of modelling the partition of gas into different phases, i.e. atomic versus molecular, which has not been attempted in most cosmological simulations of galaxy formation to date.

At the same time, there has been significant progress in understanding and modelling the formation of molecular hydrogen and SF on galactic scales. Blitz & Rosolowsky (2004, 2006) showed that the fraction of atomic to molecular gas in a sample of nearby disc galaxies was tightly correlated with the mid-plane pressure (determined by the density of both stars and gas), and this result has been confirmed in larger samples such as THINGS (Leroy et al. 2008). Robertson & Kravtsov (2008) implemented low-temperature ($T < 10^4$ K) cooling, photodissociation of H_2, and an H_2-based SF recipe in hydrodynamic simulations of isolated disc galaxies of various masses. Krumholz, McKee & Tumlinson (2009b) presented analytic models for the formation of H_2 as a function of total gas density and metallicity, supported by numerical simulations with simplified geometries (Krumholz, McKee & Tumlinson 2008, 2009a), emphasizing the importance of metallicity as a controlling parameter in H_2 formation. Gnedin & Kravtsov (2010, 2011) included detailed chemistry and low-temperature cooling as well as a simplified treatment of radiative transfer and an H_2-based SF recipe in cosmological 'zoom-in' adaptive mesh refinement simulations of small regions, and presented analytic fitting functions to their results as a function of total gas density, metallicity, and the strength of the local UV background. Christensen et al. (2012) used a similar approach to implement chemistry and simplified radiative transfer in smoothed particle hydrodynamics zoom-in simulations of galaxy-sized regions, which include a blast-wave treatment of supernova (SN) feedback.

A somewhat different view has been presented by Ostriker, McKee & Leroy (2010), who propose that heating of the interstellar medium (ISM) by the stellar UV background plays a key role in regulating SF. In their mode, the thermal pressure in the diffuse ISM, which is proportional to the UV heating rate, adjusts until it balances the mid-plane pressure set by the vertical gravitational potential. This could provide an explanation for the strong empirical correlation between H_2 fraction and disc mid-plane pressure found by Blitz & Rosolowsky (2006).

Although detailed simulations are crucial in order to understand the complex physical processes involved, extremely high resolution is required in order to obtain reliable results (see e.g. Christensen et al. 2012), implying that it will be feasible to simulate only small numbers of galaxies with these techniques for the next few years. Meanwhile, large surveys of cold gas in nearby and distant galaxies with new and upcoming facilities such as the Atacama Large Millimeter/submillimeter Array (ALMA) and the SKA (Square Kilometre Array) and its pathfinders are already being planned and pilot projects are underway. As a result, it is important to develop computationally efficient techniques that can incorporate physically motivated treatments of gas partitioning into its atomic, molecular, and possibly ionized phases and H_2-based SF recipes into simulations of cosmological volumes.

Semi-analytic models (SAMs) provide an alternative approach to this problem. In semi-analytic merger tree models, a merger tree represents the formation and growth of a dark matter halo that is identified at some redshift of interest; these merger trees may be extracted from dissipationless *N*-body simulations or created using analytic techniques (e.g. Somerville & Kolatt 1999; Parkinson, Cole & Helly 2008). Simplified but physically motivated recipes are used to track the rate of gas cooling into galaxies, and these recipes have been tested against fully numerical hydrodynamic simulations (e.g. Hirschmann et al. 2012a). These models use angular momentum based arguments to track the radial sizes of forming discs (Mo, Mao & White 1998; Somerville et al. 2008b), and can then implement recipes for how cold gas is converted into stars, and how energy and momentum from massive stars and supernovae (SNe) is returned to the ISM. This 'feedback' from stars and SNe is assumed to drive large-scale winds that can remove gas from the galaxy. The production, ejection, and recycling of metals is also tracked. Thus, our existing semi-analytic modelling framework provides the main quantities (total gas density in discs, gas metallicity) needed to implement physically motivated recipes for partitioning gas into an atomic and molecular component and then implementing an H_2-based SF recipe.

Several efforts along these lines have already been presented in the literature. Obreschkow & Rawlings (2009) implemented a prescription to estimate the H_2 fraction based on the empirical pressure-based recipe of Blitz & Rosolowsky (2006) applied in post-processing to the Millennium simulations of De Lucia & Blaizot (2007). However, in this approach, the SF in the simulations was still based on a traditional KS recipe using the total gas density, not self-consistently on the estimated H_2 gas density. Fu et al. (2010, 2012) modelled the partitioning of gas into H I and H_2 in radial bins in each galaxy, using both the metallicity-dependent recipes of Krumholz et al. (2009b) and the pressure-based recipe of Blitz & Rosolowsky (2006), and self-consistently implemented an H_2-based SF recipe, within the semi-analytic modelling framework of Guo et al. (2011). Lagos et al. (2011a,b) also estimated gas partitioning into an atomic and molecular component, and implemented an H_2-based SF recipe, within the GALFORM SAMs (Baugh et al. 2005; Bower et al. 2006). Similar modelling efforts utilizing a somewhat more simplified framework (i.e. only the mass accretion history of the largest progenitor is tracked, rather than the full merger tree) have been presented by Dutton, van den Bosch & Dekel (2010) using the pressure-based Blitz & Rosolowsky (2006) approach, and Krumholz & Dekel (2012) using the Krumholz et al. (2009b) metallicity-based approach.

It is already clear that the results of this kind of exercise may depend on the other ingredients of the modelling, in particular on the treatment of stellar feedback, chemical evolution, and potentially on feedback from active galactic nuclei (AGN). In this work, we present new models that incorporate a metallicity- or pressure-oriented treatment of atomic-molecular gas partitioning and an H_2-based SF recipe within the semi-analytic modelling framework developed by the Santa Cruz group (Somerville & Primack 1999; Somerville, Primack & Faber 2001; Somerville et al. 2008a, 2012).

The current generation of SAMs (incorporating some form of 'quenching' in massive haloes, e.g. from AGN feedback) has been fairly successful at reproducing a variety of galaxy observations, but suffer from generic problems. Both the successes and problems seem to be common to the SAMs developed by many different groups as well as to large-volume cosmological hydrodynamic simulations (see Somerville & Davé 2015, for a discussion). Significant successes include the ability to match the observed stellar mass function or luminosity functions from the UV to the NIR at $z = 0$, while simultaneously matching the gas fraction as a function of stellar mass for nearby disc galaxies (e.g. Somerville et al. 2008a, 2012; Lu et al. 2014). Observations show that massive galaxies form their stars early, and that the SF in many of these massive objects is *quenched* early, so that their stars evolve largely passively. There is some tension in the ability of models to produce enough massive galaxies at early times ($z \gtrsim 2$), and a dearth of very rapidly star-forming objects observed in the sub-mm and FIR (Niemi et al. 2012; Somerville et al. 2012). However, the evolution of the number of massive 'quenched' galaxies in models with AGN feedback seems to match observations reasonably well (Kimm et al. 2009; Brennan et al. 2015).

Low-mass galaxies seem to present a more thorny set of problems, which we refer to collectively as the 'dwarf galaxy conundrum'. Models that reproduce the low-mass end of the observed stellar mass function locally, generically *overproduce* low-mass ($m_* \lesssim 10^{10}\,\mathrm{M}_\odot$) galaxies at redshifts $0.5 \lesssim z \lesssim 2$. Moreover, low-mass galaxies apparently have (specific) star formation rates (sSFR) that are too low over the same redshift range. The stellar ages predicted by our models for these galaxies are too old compared to those derived for nearby galaxies based on 'archaeological' evidence. A summary of these problems, demonstrated for several independently developed SAMs, was presented in Fontanot et al. (2009). Weinmann et al. (2012) presented a similar study that showed that the same problems also occur in numerical hydrodynamic simulations, and recently Somerville & Davé (2015) showed that the problem persists to varying degrees in most state-of-the-art SAMs and cosmological hydrodynamical simulations. It has been suggested that these problems might be due to inaccurate recipes for SF, and that they might be cured by implementing metallicity-dependent recipes for H_2 formation and H_2-based SF (Krumholz & Dekel 2012; Kuhlen et al. 2012). This was one of the original motivations for the work we present here.

In the meantime, several works have suggested that modifying the treatment of SF physics is not the most promising solution to the low-mass galaxy problem. Fu et al. (2012) investigated different SF and gas partitioning recipes, and showed that this did not cure the overabundance of low-mass galaxies in their models. Henriques et al. (2013) and Henriques et al. (2015) showed that the problem could be largely solved by modifying the recipes for stellar-driven outflows and the re-accretion of gas ejected by these stellar driven winds in their SAMs. White, Somerville & Ferguson (2015) conducted a more empirical study and found that a wide range of variations in the star formation efficiency (SFE) could not solve the low-mass galaxy problem, while again they found that modifying the recipes for stellar-driven outflows and re-infall of ejected gas appeared most promising. In this paper, we present a simple explanation for why and how changing these different recipes in SAMs affects galaxies at different mass scales and different cosmic epochs, providing greater insight into the results of the more detailed models.

The purpose of this paper is to present the details of how we incorporate partitioning of gas into an atomic, molecular and (optionally) ionized component in our existing SAMs, how we self-consistently implement an H_2-based SF recipe, and how sensitive our results are to details of the implementation. We explore three different recipes for the partitioning of gas into different phases: the pressure-based recipe of Blitz & Rosolowsky (2006, hereafter BR) and two metallicity-based recipes, that of Krumholz et al. (KMT) and that of Gnedin & Kravtsov (2011, hereafter GK). We compare the predictions of these three new models with those using the 'classic' KS SF recipe with no gas partitioning. In addition, we explore several different empirically motivated H_2-based SF recipes.

This paper is part of a series of related works. In Popping, Somerville & Trager (2014a, hereafter PST14), we presented predictions for the atomic and molecular gas content of galaxies, and its evolution with redshift from z = 6–0, using the same models presented here. Popping et al. (2014b) extended these models by carrying out radiative transfer calculations to predict sub-mm line emission luminosities from several atomic and molecular species, including CO, HCN, C^+, and [O I]. In Berry et al. (2014), we presented predictions for the properties of objects that would be selected as Damped Lyman α systems in absorption against background quasars, again using the same model framework described here. In this paper, we focus on quantities pertaining to the stellar content, SFR, and metal content of galaxies and their evolution since $z \sim 6$. In addition, we explore a wider variety of model variants than presented in the earlier works.

The outline of the paper is as follows. In Section 2.1, we outline the basic framework of the SAMs and the treatment of structure formation, gas cooling and infall, chemical evolution, and starbursts and morphological transformation via galaxy mergers. In Section 2.2, we describe our approaches for partitioning cold gas into an atomic, molecular, and (optionally) ionized component, in Section 2.3 we describe the new H_2-based SF recipes, and in Section 2.4 we describe our implementation of metal-enhanced winds (MEW). In Section 2.5, we describe how we choose the values of the free parameters in our models, and summarize their values. In Section 3.1, we show how the SF histories and build-up of stars, gas, and metals as a function of halo mass are impacted by the different recipes for gas partitioning and SF, and other details of our model implementation. In Section 3.2, we show predictions for the relationship between total gas density and SFR density in our models. In Section 3.3, we present predictions for the stellar mass functions and stellar fractions, sSFR, gas depletion time-scales, and gas and stellar phase metallicities over cosmic time from $z \sim 6$ to the present. We discuss our results in Section 4 and summarize and conclude in Section 5.

2 MODEL DESCRIPTION

The SAMs used here have been described in detail in Somerville & Primack (1999), Somerville et al. (2001) and most recently in Somerville et al. (2008a, hereafter S08) and Somerville et al. (2012, hereafter S12). The Santa Cruz modelling framework has also recently been described in Porter et al. (2014). We refer the reader to those papers for details.

2.1 The SAM framework

This section describes the aspects of the SAMs that have been documented in previous papers. Therefore, we give a relatively brief description of these ingredients here.

In this work, the merging histories (or merger trees) of dark matter haloes are constructed based on the extended Press–Schechter (EPS) formalism using the method described in Somerville & Kolatt (1999), with improvements described in S08. These merger trees record the growth of dark matter haloes via merging and accretion, with each 'branch' representing a merger of two or more haloes. We follow each branch back in time to a minimum progenitor mass $M_{\rm res}$, which we refer to as the mass resolution of our simulation. Our SAMs give nearly identical results when run on the EPS merger trees or on merger trees extracted from dissipationless N-body simulations (Lu et al. 2014; Porter et al. 2014). We use EPS merger trees here because they allow us to attain extremely high resolution, which is important for this study. We resolve haloes down to $M_{\rm res} = 10^{10}\,{\rm M}_{\odot}$ for all root haloes, and below root halo masses of $M_{\rm res} = 10^{10}\,{\rm M}_{\odot}$, we set $M_{\rm res} = 0.01\,M_{\rm root}$, where $M_{\rm root}$ is the mass of the root halo (see Appendix A for tests supporting these choices and showing the dependence of the model predictions on the adopted mass resolution). Our root haloes cover a range from $M_{\rm h} = 5 \times 10^{8}\,{\rm M}_{\odot}$ to $5 \times 10^{14}\,{\rm M}_{\odot}$.

When dark matter haloes merge, the central galaxy of the largest progenitor becomes the new central galaxy, and all others become 'satellites'. Satellite galaxies lose angular momentum due to dynamical friction as they orbit and may eventually merge with the central galaxy. To estimate this merger time-scale we use a variant of the Chandrasekhar formula from Boylan-Kolchin, Ma & Quataert (2008). Tidal stripping and destruction of satellites are also modelled as described in S08.

Before the Universe is reionized, each halo contains a mass of hot gas equal to the universal baryon fraction times the virial mass of the halo. After reionization, the photoionizing background suppresses the collapse of gas into low-mass haloes. We use the fitting functions provided by Gnedin (2000) and Kravtsov, Gnedin & Klypin (2004), based on their hydrodynamic simulations, to model the fraction of baryons that can collapse into haloes of a given mass after reionization, and assume that the universe was fully reionized by $z = 11$, consistent with the recent analysis by the Planck survey team (Planck Collaboration XIII 2015).

When a dark matter halo collapses, or experiences a merger that at least doubles the mass of the largest progenitor, the hot gas is assumed to be shock-heated to the virial temperature of the new halo. This radiating gas then gradually cools and collapses. The cooling rate is estimated using a simple spherically symmetric model similar to the one originally suggested by White & Frenk (1991). When this model predicts a cooling radius that is greater than the virial radius, we assume that gas cools and is accreted on to the galaxy on a dynamical time. Details are provided in S08.

We assume here that the cold gas is accreted only by the central galaxy of the halo, although in reality satellite galaxies probably also continue to accrete some cold gas after they cross the virial radius of their host. In addition, we assume that all newly cooling gas initially collapses to form a rotationally supported disc. The scale radius of the disc is computed based on the initial angular momentum of the gas and the halo profile, assuming that angular momentum is conserved and that the self-gravity of the collapsing baryons causes contraction of the matter in the inner part of the halo (Blumenthal et al. 1986; Flores et al. 1993; Mo et al. 1998). This approach has been shown to reproduce the observed size versus stellar mass relation for disc-dominated galaxies from $z \sim 0$–2 (Somerville et al. 2008b). In PST14, we also showed that our models reproduce the sizes of H I discs in nearby galaxies, and sizes of CO discs out to $z \sim 2$.

SF occurs in two modes, a normal 'disc' mode in isolated discs, and a merger-driven 'starburst' mode. SF in the disc mode is modelled as described in Section 2.3 below. The efficiency and time-scale of the merger-driven burst mode is modelled as described in S08 and is a function of the merger mass ratio and the gas fractions of the progenitors. The treatment of merger-driven bursts is based on the results of hydrodynamic simulations of binary galaxy mergers (Robertson et al. 2006; Hopkins et al. 2009b).

Some of the energy from SNe and massive stars is assumed to be deposited in the ISM, resulting in the driving of a large-scale outflow of cold gas from the galaxy. The mass outflow rate is

$$\dot{m}_{\rm out} = \epsilon_{\rm SN} \left(\frac{V_0}{V_{\rm c}}\right)^{\alpha_{\rm rh}} \dot{m}_*, \qquad (2)$$

where $V_{\rm c}$ is the maximum circular velocity of the galaxy (here approximated by $V_{\rm max}$ of the dark matter halo), $\dot{m}_*$ is the SFR, $\epsilon_{\rm SN}$ and $\alpha_{\rm SN}$ are free parameters, and $V_0 = 200$ km s^{-1} is an arbitrary normalization constant. Some fraction of this ejected gas escapes from the potential of the dark matter halo, while some is deposited in the hot gas reservoir within the halo, where it becomes eligible to cool again. The fraction of gas that is ejected from the disc but retained in the halo, versus ejected from the disc and halo, is a function of the halo circular velocity (see S08 for details), such that low-mass haloes lose a larger fraction of their gas.

The gas that is ejected from the halo is kept in a larger reservoir, along with the gas that has been prevented from falling in due to the photoionizing background. This gas is assumed to accrete on to the halo on a time-scale that is proportional to the halo dynamical time (see S08 for details).

Each generation of stars produces heavy elements, and chemical enrichment is modelled in a simplified manner using the instantaneous recycling approximation. For each parcel of new stars dm_*, we also create a mass of metals $dM_Z = y\,dm_*$, which we assume to be instantaneously mixed with the cold gas in the disc. The yield y is assumed to be constant, and is treated as a free parameter. When gas is removed from the disc by SN-driven winds as described above, a corresponding proportion of metals is also removed and deposited either in the hot gas or outside the halo, following the same proportions as the ejected gas. Ejected metals also 're-accrete' into the halo along with the ejected gas, as described above.

Mergers are assumed to remove angular momentum from the disc stars and to build up a spheriod. The efficiency of disc destruction and spheroid growth is a function of progenitor gas fraction and merger mass ratio, and is parametrized based on hydrodynamic simulations of disc–disc mergers (Hopkins et al. 2009b). These simulations indicate that more 'major' (closer to equal mass ratio) and more gas-poor mergers are more efficient at removing angular momentum, destroying discs, and building spheroids. Note that the treatment of spheroid formation in mergers used here has been updated relative to S08 as described in Hopkins et al. (2009a) and Porter et al. (2014). We do not include a disc instability driven mode for spheroid growth in the models presented here.

In addition, mergers drive gas into galactic nuclei, fuelling black hole growth. Every galaxy is born with a small 'seed' black hole (BH, here we adopt $M_{\rm seed} \sim 1.0 \times 10^{4}\,{\rm M}_{\odot}$). Following a merger, any pre-existing BH are assumed to merge immediately, and the resulting hole grows at its Eddington rate until the energy being deposited into the ISM in the central region of the galaxy is sufficient to significantly offset and eventually halt accretion via a pressure-driven outflow. This results in self-regulated accretion that leaves behind BH that naturally obey the observed correlation between BH mass and spheroid mass or velocity dispersion. Our models produce good agreement with the observed luminosity function of X-ray/optical/IR detected quasars and AGN (Hirschmann et al. 2012b).

A second mode of BH growth, termed 'radio mode', is associated with powerful jets observed at radio frequencies (Croton et al. 2006; S08). In contrast to the merger-triggered, radiatively efficient mode of BH growth described above (sometimes called 'bright mode' or 'quasar mode'), in which the BH accretion is fuelled by cold gas in the nucleus, here, hot halo gas is assumed to be accreted according to the Bondi–Hoyle model (Bondi 1952). This leads to accretion rates that are typically $\lesssim 10^{-3}$ times the Eddington rate, so that most of the BH's mass is acquired during episodes of 'bright mode' accretion. However, the radio jets are assumed to couple very efficiently with the hot halo gas, and to provide a heating term that can partially or completely offset cooling during the 'hot flow' mode (we assume that the jets cannot couple efficiently to the cold, dense gas in the infall-limited or cold flow regime).

2.2 Multiphase gas partitioning

In this section, we describe in detail the updates to the model ingredients that are explored in this paper. These include partitioning of the cold gas in galactic discs into an ionized (H II), atomic (H I), and molecular (H_2) component, an option to include MEW, and a set of new H_2-based SF recipes.

At each timestep, we compute the scale radius of the cold gas disc using the angular momentum based approach described above, and assume that the *total* (H I +H_2 +H II) radial cold gas distribution is described by an exponential with scale radius $r_{\rm gas}$. We do not attempt to track the scale radius of the stellar disc separately, but make the simple assumption that $r_* = \chi_{\rm gas} r_{\rm gas}$, with $\chi_{\rm gas}$ fixed to match observed stellar scale lengths at $z = 0$. Bigiel & Blitz (2012) showed that this is a fairly good representation, on average, of the discs of nearby spirals (see also Kravtsov 2013). We then divide the gas disc into radial annuli and compute the fraction of molecular gas, $f_{\rm H_2}(r) \equiv \Sigma_{\rm H_2}(r)/[\Sigma_{\rm H_2}(r) + \Sigma_{\rm H\,I}(r)]$, in each annulus, as described below. We use a fifth-order Runga–Kutta integration scheme to compute the integrated mass of H I and H_2 in the disc, and the integrated SFR, at each timestep.

2.2.1 *Ionized gas associated with galaxies*

Most previous SAMs have neglected the ionized gas associated with galaxies, which may be ionized either by an external background or by the radiation field from stars within the galaxy. Here we investigate a simple analytic model motivated by the work of Gnedin (2012). We assume that some fraction of the total cold gas in the galaxy, $f_{\rm ion,int}$, is ionized by the galaxy's own stars. In addition, a slab of gas on each side of the disc is ionized by the external background radiation field. Assuming that all gas with a surface density below some critical value $\Sigma_{\rm H\,II}$ is ionized, we have

$$f_{\rm ion} = \frac{\Sigma_{\rm H\,II}}{\Sigma_0}\left[1 + \ln\left(\frac{\Sigma_0}{\Sigma_{\rm H\,II}}\right) + 0.5\left(\ln\left(\frac{\Sigma_0}{\Sigma_{\rm H\,II}}\right)\right)^2\right],$$

where $\Sigma_0 \equiv m_{\rm cold}/(2\pi r_{\rm gas})^2$ is the central surface density of the cold gas ($m_{\rm cold}$ is the mass of all cold gas in the disc and $r_{\rm gas}$ is the scale radius of the gas disc). We typically assume $f_{\rm ion,\,int} = 0.2$ (as in the Milky Way) and $\Sigma_{\rm H\,II} = 0.4\,{\rm M}_{\odot}\,{\rm pc}^{-2}$, as in Gnedin (2012). Applying this model within our SAM gives remarkably good agreement with the ionized fractions as a function of circular velocity shown in fig. 2 of Gnedin (2012), obtained from hydrodynamic simulations with time dependent and spatially variable 3D radiative transfer of ionizing radiation from local sources and the cosmic background.

2.2.2 *Molecular gas: pressure-based partitioning*

We consider three approaches for computing the molecular gas fractions in galaxies. The first is based on the empirical pressure-based recipe presented by BR, and will be referred to as the BR recipe. BR found a power-law relation between the disc mid-plane pressure and the ratio between molecular and atomic hydrogen, i.e.

$$R_{\rm H_2} = \left(\frac{\Sigma_{\rm H_2}}{\Sigma_{\rm H\,I}}\right) = \left(\frac{P_{\rm m}}{P_0}\right)^{\alpha_{\rm BR}}, \quad (3)$$

where $\Sigma_{\rm H_2}$ and $\Sigma_{\rm H\,I}$ are the H_2 and H I surface density, P_0 and $\alpha_{\rm BR}$ are free parameters that are obtained from a fit to the observational data, and $P_{\rm m}$ is the mid-plane pressure acting on the galactic disc. We adopted $\log P_0/k_B = 4.23$ cm^3 K and $\alpha_{\rm BR} = 0.8$, based on observations from Leroy et al. (2008).

We estimate the hydrostatic pressure as a function of the distance from the centre of the disc r as

$$P(r) = \frac{\pi}{2} G\Sigma_{\rm gas}(r)[\Sigma_{\rm gas}(r) + f_\sigma(r)\Sigma_*(r)], \quad (4)$$

where G is the gravitational constant, $\Sigma_{\rm gas}$ is the cold gas surface density, Σ_* is the stellar surface density, and f_σ is the ratio of the vertical velocity dispersions of the gas and stars: $f_\sigma(r) = \frac{\sigma_{\rm gas}}{\sigma_*}$. Following Fu et al. (2010), we adopt $f_\sigma(r) = 0.1\sqrt{\Sigma_{*,0}/\Sigma_*}$, where $\Sigma_{*,0} \equiv m_*/(2\pi r_*^2)$, based on empirical scalings for nearby disc galaxies.

2.2.3 *Molecular gas: metallicity-based partitioning*

Gnedin and Kravtsov (Gnedin & Kravtsov 2010; GK) performed high-resolution 'zoom-in' cosmological simulations with the adaptive refinement tree code of Kravtsov, Klypin & Khokhlov (1997), including gravity, hydrodynamics, non-equilibrium chemistry, and simplified on-the-fly radiative transfer. These simulations are therefore able to follow the formation of molecular hydrogen through primordial channels and on dust grains, as well as dissociation of molecular hydrogen and self- and dust-shielding.

GK presented a fitting function based on their simulations, which effectively parametrizes the fraction of molecular hydrogen as a function of the dust-to-gas ratio relative to the Milky Way, $D_{\rm MW}$, the UV ionizing background relative to the Milky Way $U_{\rm MW}$, and the neutral gas surface density $\Sigma_{\rm H\,I+H_2}$. The fraction of molecular hydrogen is given by

$$f_{\rm H_2} = \left[1 + \frac{\tilde{\Sigma}_*}{\Sigma_{\rm H\,I+H_2}}\right]^{-2},$$

where

$$\tilde{\Sigma}_* = 20\,{\rm M}_{\odot}{\rm pc}^{-2}\frac{\Lambda^{4/7}}{D_{\rm MW}}\frac{1}{\sqrt{1 + U_{\rm MW}D_{\rm MW}^2}}$$

$$\Lambda = \ln(1 + gD_{\rm MW}^{3/7}(U_{\rm MW}/15)^{4/7})$$

$$g = \frac{1+\alpha s + s^2}{1+s}$$
$$s = \frac{0.04}{D_* + D_{\rm MW}}$$
$$\alpha = 5\frac{U_{\rm MW}/2}{1+(U_{\rm MW}/2)^2}$$
$$D_* = 1.5\times10^{-3}\,\ln(1+(3U_{\rm MW})^{1.7}).$$

We take the dust-to-gas ratio to be equal to the cold gas phase metallicity in solar units $D_{\rm MW} = Z/Z_\odot$. The local UV background relative to the MW is assumed to scale in proportion with the global SFR of the galaxy in the previous time step relative to the MW SFR, $U_{\rm MW} = \frac{\rm SFR}{\rm SFR_{MW}}$, where we choose $\mathrm{SFR_{MW}} = 1.0\,\mathrm{M_\odot\,yr^{-1}}$ (Murray & Rahman 2010; Robitaille & Whitney 2010). We refer to this as our 'fiducial' GK model. We also investigate the results of keeping $U_{\rm MW}$ fixed to the Milky Way value, which we refer to as the GKFUV model (fixed UV field).

An alternate approach based on similar physical processes was presented in a series of papers by Krumholz and collaborators (Krumholz et al. 2008, 2009a,b). Krumholz et al. (2009b) developed an analytic model for the molecular fraction in galaxies, based on the ansatz that the interplay between the interstellar radiation field and molecular self-shielding determines the molecular fraction. They presented a fitting function:

$$f_{\rm H_2} = 1 - \left[1+\left(\frac{3}{4}\frac{s}{1+\delta}\right)^{-5}\right]^{-1/5},$$

where $s = \ln(1+0.6\chi)/(0.04\Sigma_{\rm comp,\,0}Z')$, $\chi = 0.77(1+3.1Z'^{0.365})$, $\delta = 0.0712(0.1s^{-1}+0.675)^{-2.8}$, $\Sigma_{\rm comp,\,0} = \Sigma_{\rm comp}/(1\mathrm{M_\odot\,pc^{-2}})$, and $Z' \equiv Z/Z_\odot$. The quantity denoted $\Sigma_{\rm comp}$ is the surface density within an ∼100 pc sized atomic-molecular cloud complex. Krumholz et al. (2009b) suggest using a 'clumping factor' to apply the model to simulations with spatial resolution coarser than 100 pc; i.e. $\Sigma_{\rm comp} \to c\Sigma_{\rm H\,I+H_2}$, where $c \geq 1$, and $\Sigma_{\rm H\,I+H_2}$ is the neutral gas surface density on some larger scale. The appropriate value of c depends on this spatial scale, where $c \to 1$ as the scale approaches 100 pc. We refer to this as the KMT gas partitioning recipe.

Both GK and Krumholz et al. (2009b) note that the fitting functions, as well as perhaps (in the case of KMT) the underlying assumptions of the model, begin to break down at metallicities lower than about 1/50th of the Solar value.

The KMT and GK fitting functions above characterize the formation of H_2 on dust grains, which is the dominant mechanism once the gas is enriched to more than a few hundredths of Solar metallicity. Other channels for the formation of H_2 in primordial gas must be responsible for producing the molecular hydrogen out of which the first stars were formed. Hydrodynamic simulations containing detailed chemical networks and analytic calculations have shown that H_2 can form in metal-free gas in dark matter haloes above a critical mass $M_{\rm crit} \sim 10^5\,\mathrm{M_\odot}$ (e.g. Nakamura & Umemura 2001; Glover 2013). This gas can then form 'Pop III' stars which can enrich the surrounding ISM to $Z_{\rm III} \sim 10^{-3}\,\mathrm{Z_\odot}$ (Schneider et al. 2002; Greif et al. 2010; Wise et al. 2012). These processes take place in haloes much smaller than our resolution limit. We represent them by setting a 'floor' to the molecular hydrogen fraction in our haloes, $f_{\rm H_2,floor}$. In addition, we 'pre-enrich' the initial hot gas in haloes, and the gas accreted on to haloes due to cosmological infall, to a metallicity of $Z_{\rm pre-enrich}$. We adopt typical values of $f_{\rm H_2,floor} = 10^{-4}$ and $Z_{\rm pre-enrich} = 10^{-3}\,\mathrm{Z_\odot}$ (Haiman, Rees & Loeb 1996; Bromm & Larson 2004). We discuss the sensitivity of our results to these parameters in Appendix A. Note that observations of resolved stars in the halo of our Galaxy and local dwarfs have revealed stars with metallicities below $Z \sim 10^{-3}\,\mathrm{Z_\odot}$ (Schörck et al. 2009; Tolstoy, Hill & Tosi 2009; Kirby et al. 2011), precluding much higher values for $Z_{\rm pre-enrich}$.

2.3 SF recipes

2.3.1 The KS recipe

The KS recipe (Kennicutt 1998) assumes that the surface density of SF in a galaxy is a function of the *total* surface density of the cold neutral gas (atomic and molecular), above some threshold surface density $\Sigma_{\rm crit}$.

The star formation rate density (SFRD, per unit area) for $\Sigma_{\rm gas} > \Sigma_{\rm crit}$ is given by

$$\Sigma_{\rm SFR} = A_{\rm SF}\,\Sigma_{\rm gas}{}^{N_{\rm SF}}, \tag{5}$$

where $\Sigma_{\rm SFR} = 0$ for $\Sigma_{\rm gas} < \Sigma_{\rm crit}$. This recipe is the same one used in most of our previously published SAMs (S08; S12), and is similar to recipes commonly adopted in many other SAMs. The values of the free parameters are given in Table 1.

2.3.2 Molecular hydrogen-based recipes

In the same spirit as the KS recipe, we use empirical relationships from observations to motivate our H_2-based recipes. Bigiel et al. (2008) found, based on observations of spiral galaxies from the THINGS survey, that the SFR surface density can be directly related to the surface density of molecular gas, i.e.

$$\Sigma_{\rm SFR} = \left(\frac{A_{\rm SF}}{10\,\mathrm{M_\odot\,pc^{-2}}}\right)\,\Sigma_{\rm H_2}{}^{N_{\rm SF}} \tag{6}$$

with $N_{\rm SF} \simeq 1$ (see also Bigiel et al. 2011; Leroy et al. 2013). Observations of higher density environments suggest that above some critical H_2 surface density, the slope of the relation described in equation (6) steepens (Narayanan et al. 2012). We therefore also consider a two-part scaling law given by

$$\Sigma_{\rm SFR} = A_{\rm SF}\left(\frac{\Sigma_{\rm H_2}}{10\,\mathrm{M_\odot\,pc^{-2}}}\right)\left(1+\frac{\Sigma_{\rm H_2}}{\Sigma_{\rm H_2,crit}}\right)^{N_{\rm SF}}. \tag{7}$$

The values of the parameters $A_{\rm SF}$, $N_{\rm SF}$, and $\Sigma_{\rm H_2,crit}$ are given in Table 1.

An SF relation that changes slope above a critical density is also expected based on theoretical grounds. Krumholz et al. (2009b) adopt the SF relation:

$$\Sigma_{\rm SFR} = A_{\rm SF}\Sigma_{\rm H_2}\left(\frac{\Sigma_{\rm gas}}{\Sigma_{\rm crit}}\right)^{N_{\rm SF}}, \tag{8}$$

where $N_{\rm SF} = -0.33$ for $\Sigma_{\rm gas}/\Sigma_{\rm crit} < 1$ and $N_{\rm SF} = 0.33$ for $\Sigma_{\rm gas}/\Sigma_{\rm crit} > 1$. KMT adopt $A_{\rm SF} = 1/2.6\ \mathrm{Gyr^{-1}}$ and $\Sigma_{\rm crit} = 85\,\mathrm{M_\odot pc^{-2}}$. We adopt the same parameter values in our 'KMT' model.

A comparison of various SF relations in the literature, and used in this work, is shown in Fig. 1. An extensive discussion of how our adopted SF recipes compare with those used in the literature may be found in Section 4.4.

2.4 Metal-enhanced winds

Most of our previous models have assumed that metals are ejected from galaxies with the same efficiency as the gas, i.e. with the

Table 1. Summary of model parameters.

Parameter	Description	Section defined	Value
$\epsilon_{\rm SN}$	SN feedback efficiency	2.1	1.5
$\alpha_{\rm SN}$	SN feedback slope	2.1	−2.2
y	Chemical yield	2.1	1.6
$\kappa_{\rm AGN}$	Radio mode AGN feedback	S08 Section 2.11, equation (20)	3.8×10^{-3}
$\chi_{\rm gas}$	Ratio of stellar to gas scale length	2.2	0.59
$\Sigma_{\rm H\,II}$	Critical density for ionized gas	2.2.1	$0.4\ {\rm M_{\odot}\ pc^{-2}}$
$f_{\rm ion,\ int}$	Internal fraction of ionized gas	2.2.1	0.2
P_0	Pressure scaling in BR recipe	2.2.2	$4.23\ k_B\ {\rm cm^3\ K}$
$\alpha_{\rm BR}$	Slope in BR recipe	2.2.2	0.8
c	Clumping factor in KMT recipe	2.2.3	5
$f_{\rm H_2,floor}$	Primordial H_2 fraction	2.2.3	10^{-4}
$Z_{\rm pre-enrich}$	Metallicity due to Pop III stars	2.2.3	10^{-3}
$\zeta_{\rm lo}$	MEW normalization	2.4	0.1
$M_{\rm ret}$	MEW mass scale	2.4	0.9
$A_{\rm SF}$ (KS)	SF relation normalization	2.3, equation (5)	1.1×10^{-4}
$N_{\rm SF}$ (KS)	SF relation slope	2.3, equation (5)	1.4
$\Sigma_{\rm crit}$ (KS)	Critical density for SF	2.3, equation (5)	$6\ {\rm M_{\odot}\ pc^{-2}}$
$A_{\rm SF}$ (Big1, Big2)	SF relation normalization	2.3, equation (6) and (7)	4.0×10^{-3}
$N_{\rm SF}$ (Big1, Big2)	SF relation slope	2.3, equation (6) and (7)	1.0
$\Sigma_{\rm H_2,crit}$	Critical H_2 density	2.3, equation (7)	$70\ {\rm M_{\odot}\ pc^{-2}}$

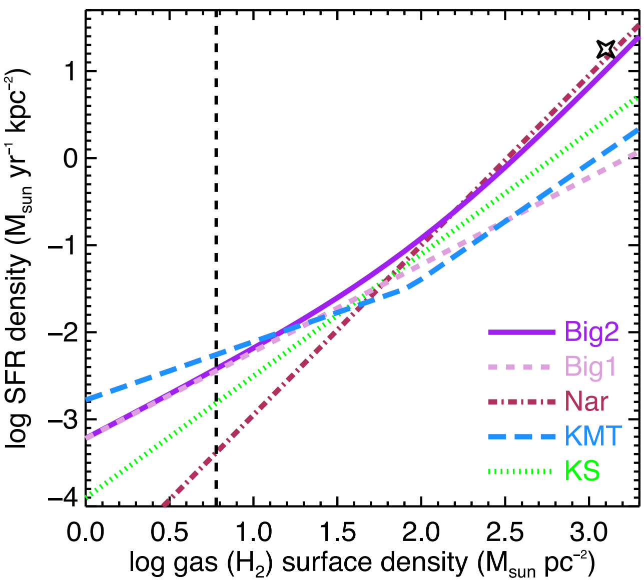

Figure 1. Empirically based SF recipes used as input in our models, and from the literature. The solid purple line shows the two-slope H_2-based recipe (Big2); the dashed lavender line shows the single slope H_2-based recipe (Big1); the dot–dashed dark red line shows the recipe based on the analysis of Narayanan et al. (2012); the dotted green line shows the (H I + H_2) based 'classic' KS recipe. The vertical dashed line shows the critical total gas surface density used in our models that implement the classic KS recipe ($\Sigma_{\rm crit}$). The star symbol shows the relation derived by Sharon et al. (2013) for an extreme starburst galaxy. Note that the Narayanan et al. (2012) results are shown for reference only; we do not show the results of incorporating this recipe in our models here.

same mass loading factor $\eta \equiv \dot{m}_{\rm out}/\dot{m}_*$. However, since metals are produced by the same massive stars and SNe that are believed to drive galactic outflows, it is possible that metals are preferentially ejected (i.e. have a higher effective mass loading factor than the gas averaged over the whole disc). Because two of our recipes for gas partitioning depend on the gas metallicity, the dispersal of metals in our models has a potentially important impact on our results. We therefore include an optional treatment of MEW in our models.

We base our parametrization of MEW on the approach used in Krumholz & Dekel (2012), in part because we want to be able to compare our results with theirs. The fraction of metals that is ejected is parametrized by

$$\zeta = \zeta_{\rm lo} \exp(-M_{\rm h}/M_{\rm ret}),$$

where both $\zeta_{\rm lo}$ and $M_{\rm ret}$ are free parameters, and $M_{\rm h}$ is the virial mass of the halo. The modified equation for the evolution of the mass in metals in the cold gas phase is then

$$\dot{M}_Z = y(1-R)(1-\zeta)\dot{m}_* + Z_{\rm hot}\dot{m}_{\rm inf} - Z_{\rm cold}\dot{m}_{\rm out},$$

where R is the recycled fraction, y is the chemical yield, $Z_{\rm hot}$ is the metallicity of the hot gas, $Z_{\rm cold}$ is the metallicity of the cold gas, and $\dot{m}_*$, $\dot{m}_{\rm inf}$, and $\dot{m}_{\rm out}$ are the SFR, inflow rate of cold gas into the disc, and the outflow rate of gas from the disc, respectively.

2.5 Calibrating the free parameters

As in any cosmological simulation, we must parametrize the sub-grid physics in our models. In keeping with common practice, we choose the values of the free parameters by tuning to a subset of observations in the local universe. In this subsection, we summarize the values of the free parameters used here (see Table 1) and the observations we used to constrain them. The parameters are the same as those used in the models presented in PST14. The values of any parameters not specified in Table 1 may be found in S08 and are the same as the ones adopted there.

We assume values for the cosmological parameters consistent with the five year *WMAP* results (*WMAP*5): $\Omega_{\rm m} = 0.28$, $\Omega_\Lambda = 0.72$, $H_0 = 70.0$, $\sigma_8 = 0.81$, and $n_{\rm s} = 0.96$ (Komatsu et al. 2009). We note that these values are generally consistent with those obtained from the analysis of the seven-year *WMAP* data release (Komatsu et al. 2011) and the recent Planck analysis (Planck Collaboration

XIII 2015). The adopted baryon fraction is 0.1658. We assume a recycled fraction of $R = 0.43$, as appropriate for a Chabrier stellar initial mass function (Chabrier 2003).

As discussed in S08 (see also White et al. 2015), in our models the SN feedback parameters mainly control the low-mass end of the stellar mass function ($m_* \lesssim M_{\rm char}$, where $M_{\rm char}$ is the characteristic mass of the 'knee' in the Schechter function describing the stellar mass function), or equivalently, the fraction of baryons that is turned into stars in haloes with $M_{\rm h} \lesssim 10^{12}\,{\rm M}_{\odot}$. On the other side, the efficiency of the 'radio mode' AGN feedback (one can think of this schematically as the efficiency with which radio jets couple to and heat the hot intragroup and intracluster medium) controls the number density of massive galaxies $m_* \gtrsim M_{\rm char}$, or the fraction of baryons that are able to turn into stars in massive haloes ($M_{\rm h} \gtrsim 10^{12}\,{\rm M}_{\odot}$). We tune the parameters controlling SN feedback and AGN feedback to reproduce the observed stellar mass function at $z = 0$ in our traditional KS model (see S08 for details). These parameters are then kept fixed as we explore the effects of varying the modelling of gas partitioning and SF.

The parameters of the SF recipe mainly control the fraction of cold gas in galaxies, and do not strongly affect the $z = 0$ stellar mass function (White et al. 2015). We require the parameters characterizing our SF recipes to lie within the observational uncertainties from recent empirical constraints, and tune them within these limits to match the *total* gas fractions as a function of galaxy stellar mass at $z = 0$ (see PST14).

The chemical yield y could in principle be obtained from stellar evolution models, but these model yields are uncertain by a factor of ~2, and the single-element instantaneous recycling approach to chemical evolution that we are using here is somewhat crude, so we instead treat the chemical yield as a free parameter (though we restrict it to be in the expected range). We tune our yield to match the normalization of the observed *stellar* metallicity versus mass relation of Gallazzi et al. (2005).

We take the parameter values for the BR gas partitioning recipe from the observational results of Leroy et al. (2008). We implement the GK recipe as it is given in GK, with no tunable parameters. The KMT recipe has one free parameter, the clumping factor of the gas. We adopt $c = 5$, following Krumholz & Dekel (2012).

Our new models predict the fraction of cold gas in different phases: ionized, atomic, and molecular. We showed the predictions of our two fiducial models (GK and BR) for the fraction of H I and H_2 as a function of galaxy internal stellar density and stellar mass in PST14. It is encouraging that our new models reproduce these observed scalings for nearby galaxies quite well without any additional tuning. We emphasize that the only new free parameters in the gas partitioning recipe have been taken directly from observations (in the case of the BR recipe) or from numerical simulations (in the case of KMT and GK). We summarize the ingredients of different models investigated in this paper in Table 2.

Table 2. Summary of model variants.

Model	H I/H_2 partitioning	SF law	MEW
KS 'fiducial'	None	KS	N
BR 'fiducial'	BR	Big2	N
GK 'fiducial'	GK, $U_{\rm MW} \propto$ SFR	Big2	N
GK+Big1	GK, $U_{\rm MW} \propto$ SFR	Big1	N
GKFUV	GK, $U_{\rm MW} = 1$	Big2	N
BR+Big1	BR	Big1	N
KMT+Big1	KMT	Big1	N
KMT	KMT	KMT	N
KMT+MEW	KMT	KMT	Y

Table 3. Halo mass to stellar mass translation for fiducial GK model.

$\log M_{\rm h}[{\rm M}_{\odot}]$	$\log m_*[{\rm M}_{\odot}]$
10.0	6.78
10.5	8.0
11.0	8.89
11.5	9.74
12.0	10.65
13.0	11.13
14.0	11.52
14.5	11.69

3 RESULTS

3.1 Effects of varying model ingredients, parameter values, and resolution

In this subsection, we explore the sensitivity of our model results to our new model ingredients and parameter values related to gas partitioning and SF, as well as to our numerical resolution. Readers who are mainly interested in model predictions that are directly comparable with observations may wish to skip this subsection. To illustrate these effects, we show the properties of the central largest progenitor galaxy as a function of time (or redshift) in a set of haloes with masses at $z = 0$ ranging from $\log M_{\rm h}/{\rm M}_{\odot} = 10.0$–$11.5$ (except in one case, where we show a set of more massive haloes up to $\log M_{\rm h}/{\rm M}_{\odot} = 14.5$). The variations that we explore in each set of plots produce no significant differences for galaxies in haloes outside of the mass range shown. For the convenience of the reader, we provide in Table 3 the values of the stellar mass of the central galaxy at $z = 0$ in our fiducial GK model for each of the halo mass bins we consider. We use the same merger trees for each model, and fix all model properties that are chosen from random distributions to their average values. Each panel shows the average over 60 different realizations of haloes with the specified final mass. For each experiment we show the stellar mass, total neutral cold gas mass (H I +H_2), SFR, H_2 fraction, and gas phase metallicity. To facilitate comparison, the stellar mass, gas mass, and SFR are normalized by dividing by the root halo mass at $z = 0$. In Appendix A, we show several tests that are important to document but are not central to the main results of this paper; these include changing the mass resolution of our dark matter halo merger trees, changing the value of the pre-enrichment metallicity $Z_{\rm pre-enrich}$, switching MEW on and off, and including a component of ionized gas in galactic discs or not. In this subsection, we describe the results of tests that illustrate the effects of changing the recipes for partitioning the gas into H I and H_2, and the recipes for converting H_2 into stars. These tests are highly relevant for interpreting the global results that we will present later.

A brief summary of the tests presented in Appendix A is as follows. (1) Our model results are very robust to varying the mass resolution of our merger trees, as long as haloes are resolved down to progenitor masses about 0.01 times the final root mass. (2) Our results are quite insensitive to the adopted values for the metallicity of pre-enriched gas $Z_{\rm pre-enrich}$ and the floor on the H_2 fraction $f_{\rm H_2,floor}$, within a reasonable range of values. (3) Adopting MEW leads to later enrichment in low-mass galaxies, and a steeper dependence of metallicity on stellar mass. (4) Including an 'inert' ionized

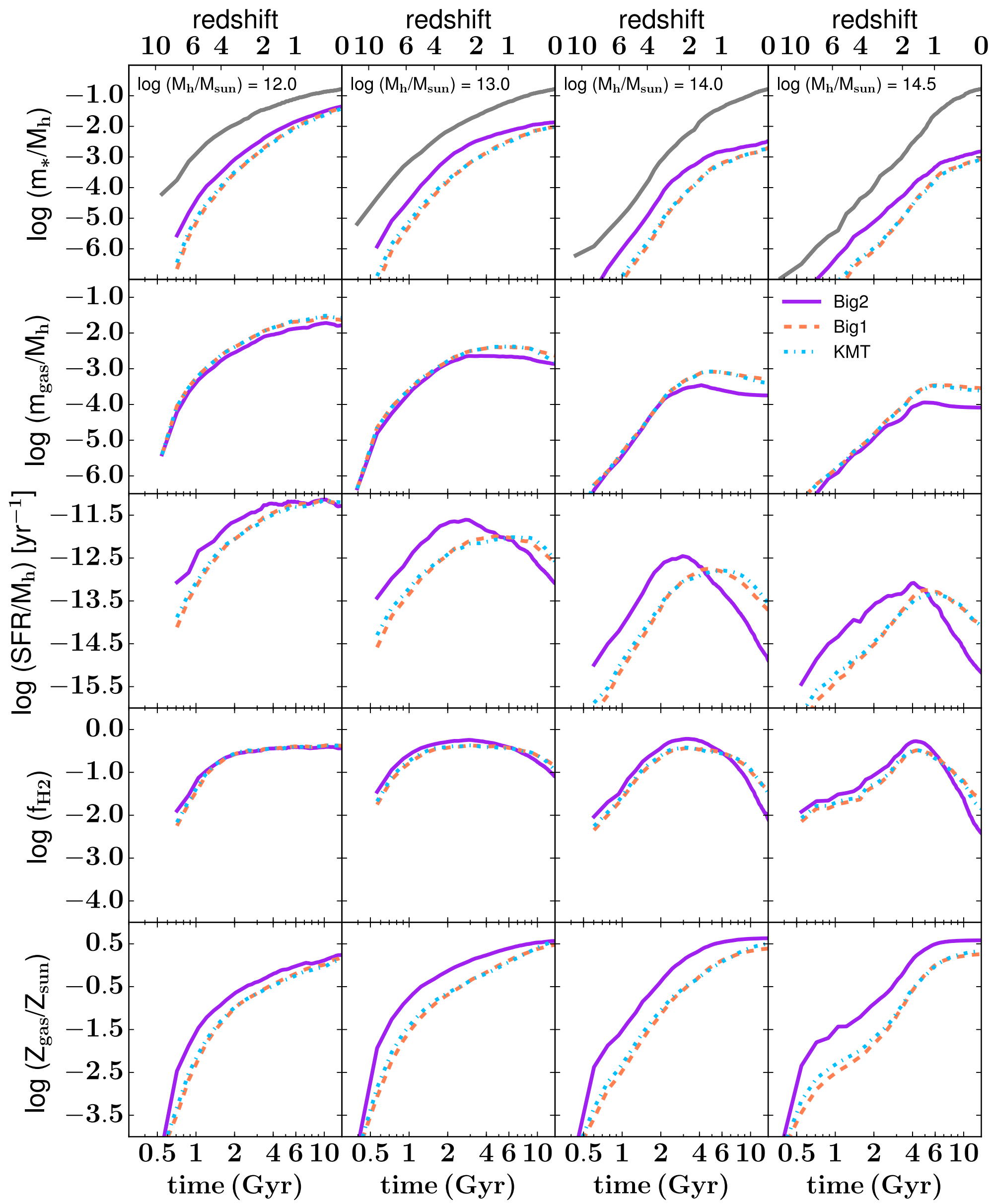

Figure 2. From top to bottom, coloured lines show the stellar mass, cold neutral gas mass (H I + H_2), and SFR normalized by the mass of the root halo at $z = 0$ for the central largest progenitor galaxy as a function of cosmic time (redshift). The H_2 fraction ($f_{H_2} \equiv m_{H_2}/(m_{H_2} + m_{H\,I})$ and gas phase metallicity (in solar units) are also shown. Grey lines show the maximum baryon fraction in the halo, $f_b\, M_h(t)$, where $M_h(t)$ is the mass of the largest progenitor halo at time t and f_b is the universal baryon fraction. We compare different recipes for converting molecular gas into stars within our fiducial GK models: Big2, Big1, and KMT. Note the different range of halo masses from the other plots. Here, the strongest effect seen is the more efficient production of stars and earlier enrichment in massive haloes in the Big2 model. This is owing to the non-linear dependence of the SF efficiency on H_2 density at high densities in this model.

gas component in our models has a negligible effect on the results presented in this paper.

In Fig. 2, we investigate several different recipes for converting cold molecular gas into stars within the fiducial GK model. We consider two variants of the empirical recipe based on the observations by Bigiel et al. (2008). In addition, we consider the recipe proposed by KMT based on theoretical arguments (see Section 2.3 for details). Big1 refers to equation (6) and Big2 to equation (7). In contrast to most of our other experiments, these variations have almost no discernable effect on the low-mass haloes ($M_h \lesssim 10^{12}\,M_\odot$). However, the Big2 recipe leads to significantly earlier build-up of stellar mass, more efficient SF at high redshift, and earlier metal enrichment in *massive* haloes ($\log M_h/M_\odot \gtrsim 12.0$). This is owing to the non-linear dependence of the SF efficiency on H_2 density at high densities in this model, and the positive correlation between halo mass and galaxy surface density in our models.

Next we experiment with changing the recipe for partitioning gas into H_2 (Fig. 3). All other ingredients are the same as our fiducial GK models. We show the pressure-based BR model as well as two alternate metallicity-based models. Recall that in the fiducial GK model, f_{H_2} depends on the total gas density and metallicity as well as the local UV radiation field (which we scale with the global galaxy SFR). In the GKFUV model, we remove the UV radiation field dependence by using the Milky Way value in all galaxies. In the KMT model, f_{H_2} depends only on total gas density and metallicity, and has different dependences on these quantities than the

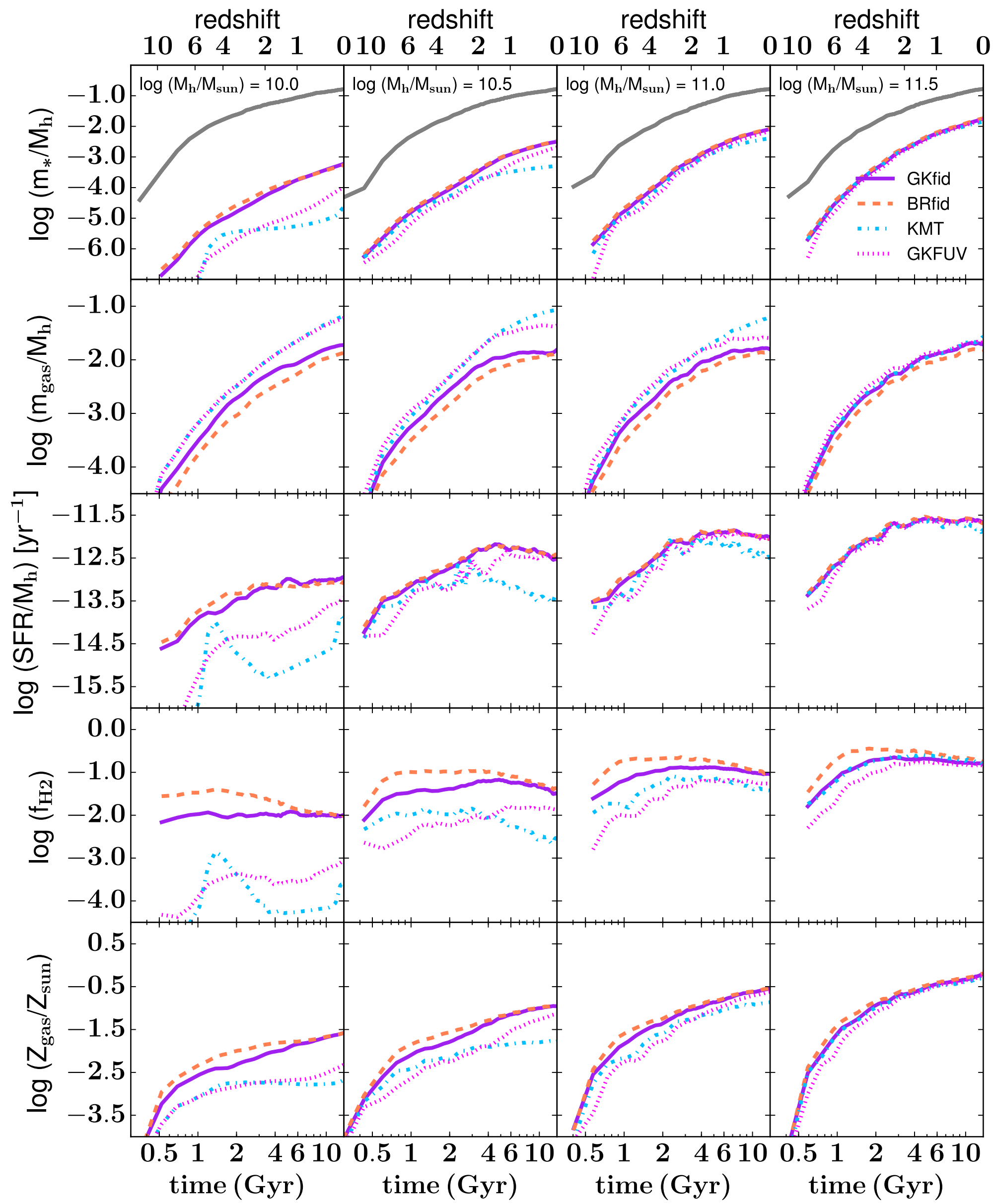

Figure 3. Same as Fig. 2, except here we compare our fiducial GK and BR models with the KMT recipe for gas partitioning and the GK recipe for gas partitioning with a fixed UV background field. The KMT recipe and the GKFUV recipes, which both neglect the dependence of H_2 formation on the local UV radiation field, both predict much lower H_2 fractions in low-mass haloes, especially at high redshift, leading to later stellar mass assembly, slightly higher overall gas fractions, lower SFE, and later metal enrichment.

GK model. The results of this experiment are quite interesting. The predictions of the BR and fiducial GK models are quite similar, although the GK model tends to predict higher gas masses, lower H_2 fractions, and lower metallicities at early times in the two lowest mass halo bins. The KMT and GKFUV models also produce similar results, as expected based on the findings of Krumholz & Gnedin (2011), who also showed the two models to be very similar. However, the KMT and GKFUV models predict significantly suppressed H_2 formation, leading to lower SFR and stellar masses, reduced chemical enrichment, and higher gas fractions in the two lowest mass halo bins. This is because galaxies in low-mass haloes tend to have lower metallicities but also lower SFR. Therefore in our fiducial GK models, the lower metallicity, which makes H_2 formation less efficient, is mitigated by the lower SFR, which leads to weaker photodissociation and relatively higher H_2 fractions.

In our penultimate experiment (Fig. 4), we compare our two new fiducial multiphase gas recipes with the ‘classic’ KS recipe (see Section 2.3), in which all cold gas above a fixed surface density is eligible for SF. This KS recipe has been used in many previous SAMs (e.g. S08; S12; Porter et al. 2014). The build-up of stellar mass is almost identical in all three models. However, the cold gas mass is highest in the GK model and tends to be lowest in the KS model. The H_2 fraction is also lower at early times in the lowest halo mass bin in the GK model. Note that the H_2 fraction shown for the KS model has been computed using the BR recipe in post-processing, but this has no impact on the SF in the model. Overall, the degree of similarity between the results of the three models is quite surprising, given the rather different physical premises on which they are based. We discuss possible reasons for this in Section 4.

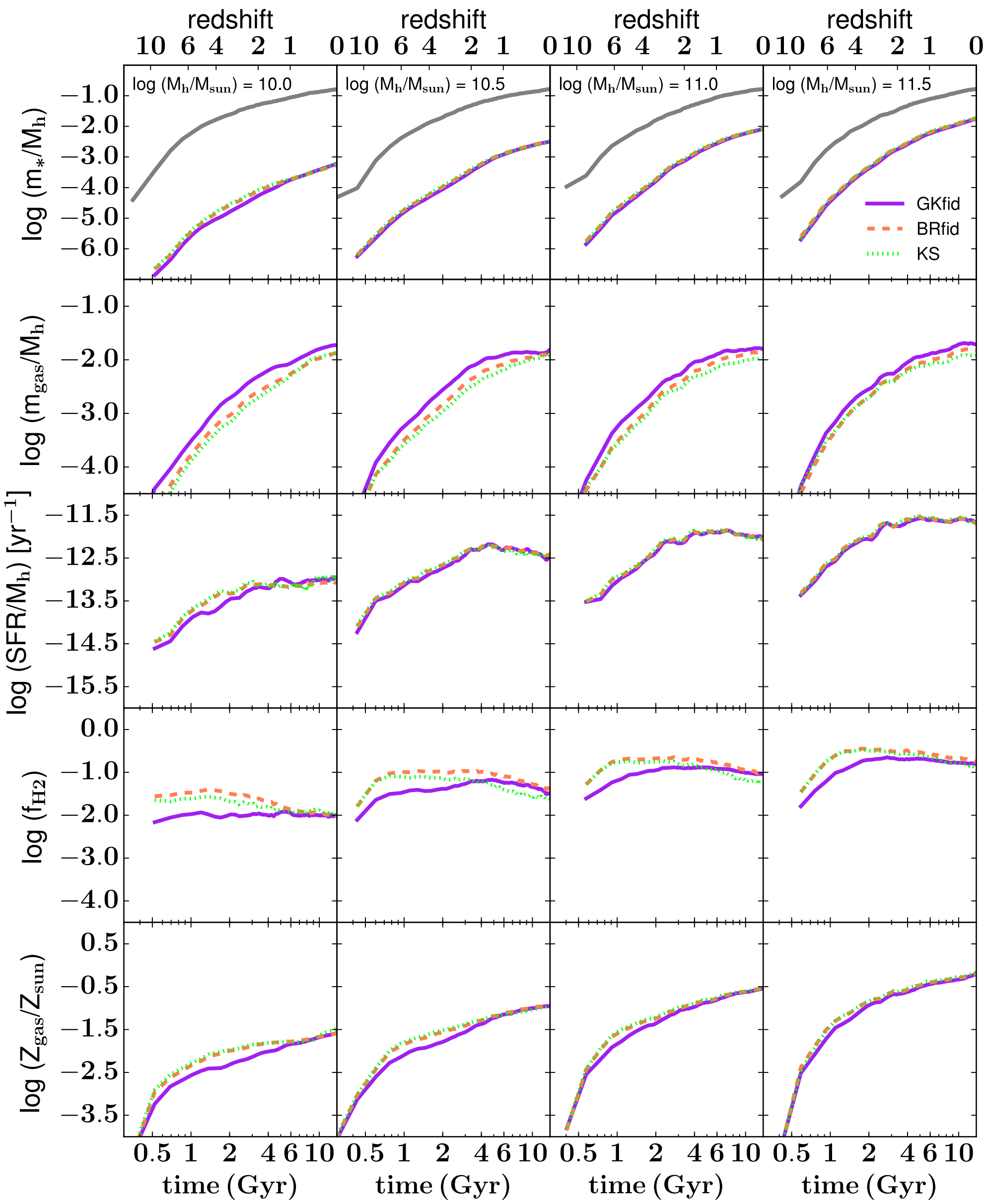

Figure 4. Same as Fig. 2, except here we compare the three 'fiducial' models, GK, BR, and KS. We find remarkably similar results among all three models. The largest differences are in the predicted overall gas fraction and the H_2 fraction at high redshift in the lowest mass haloes.

In our final experiment, shown in Fig. 5, we compare the evolution in our fiducial GK model with variants that include combinations of recipes that are similar to those used in several published models from the literature. For example, the BR+Big1 model treats gas partitioning and conversion of H_2 to stars using similar recipes to the BR model of Lagos et al. (2011b) and the 'Bigiel + H_2 prescription 2' of Fu et al. (2012). The KMT+Big1 model contains similar ingredients to the 'Krumholz + H_2 prescription 1' of Fu et al. (2012). In the KMT+MEW model, we use the KMT recipes for both gas partitioning and SF, as well as including MEW, as in the models of Krumholz & Dekel (2012). Note that the KMT model of Lagos et al. (2011b) adopts the KMT recipes for both gas partitioning and SF, but does not adopt MEW, so does not correspond exactly to any of the cases shown here. However, adopting these choices in our models yields results very similar to the KMT+Big1 model shown. We emphasize that many other aspects of our models differ from those used by other SAMs in the literature, so these may not correspond to the actual predictions of those models. This exercise is intended to shed some light on the effect of choosing different recipes for gas partitioning and conversion of H_2 into stars in a controlled environment where all other aspects of the models are held fixed. Most other SAMs to date that have attempted to track multiphase gas with a metallicity-based approach have done so using the KMT recipe. Our experiment shows that this may result in more delayed SF and enrichment in low-mass haloes than our fiducial models predict. Krumholz & Dekel (2012) additionally adopt strongly halo mass dependent MEW. These two effects together lead to strong suppression of SF in low-mass galaxies ($m_* \lesssim 10^8 M_\odot$).

3.2 SF relations

Fig. 6 shows the relationship between the total neutral cold gas surface density $\Sigma_{H_I+H_2}$ and SFRD Σ_{SFR} in our two new fiducial models with multiphase gas partitioning. In our previous generation

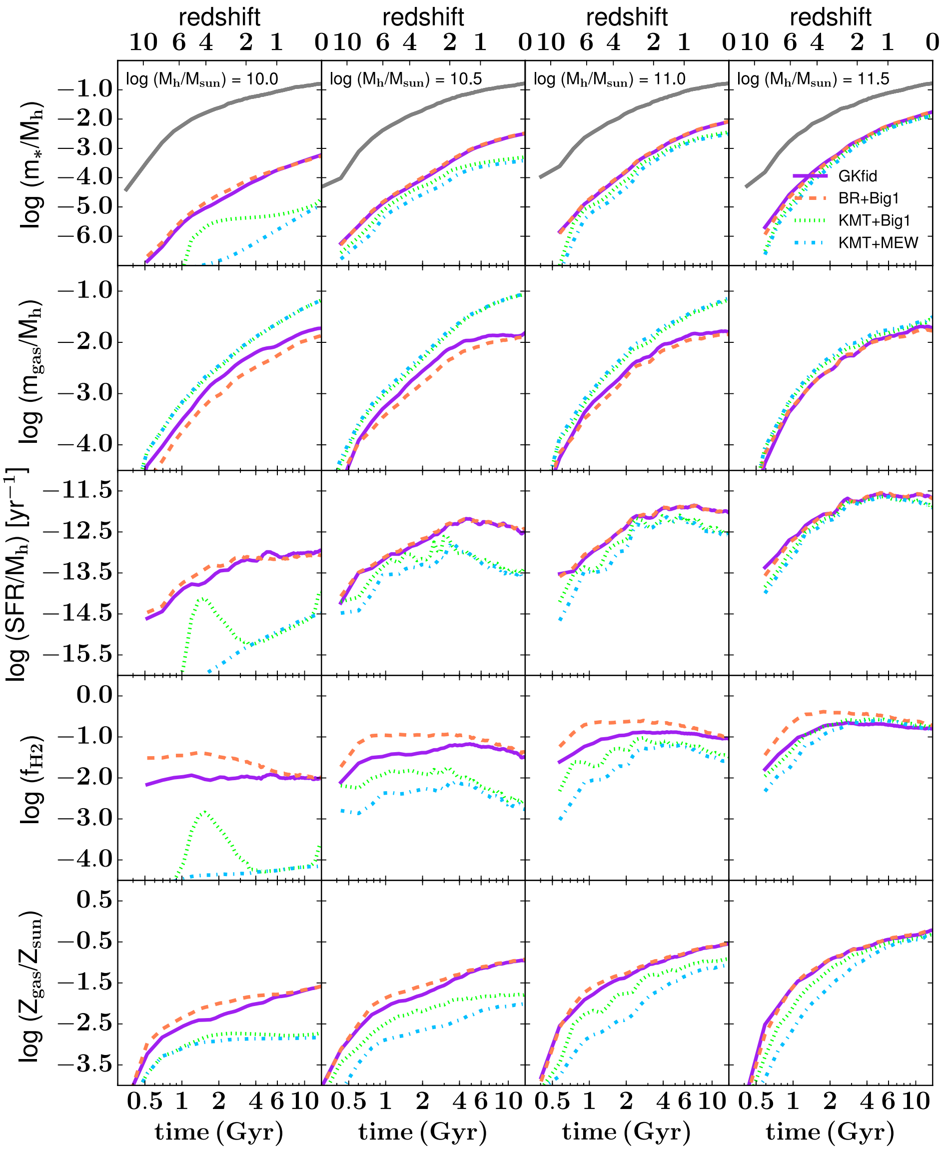

Figure 5. Same as Fig. 2, except here we compare the results of our fiducial GK model with combinations of model ingredients that are similar to those used in several models in the literature. Models that neglect the effect of a varying UV background predict later SF, higher cold gas fractions, lower H_2 fractions, and later chemical enrichment in low-mass haloes. MEW, when coupled with metallicity-dependent gas partition recipes, further delay SF and enrichment.

of models, galaxies had a deterministic relation between $\Sigma_{\mathrm{H\,I+H_2}}$ and Σ_{SFR} given by the assumed KS relation (as plotted in Fig. 1), and Σ_{SFR} was set to zero below the critical gas surface density Σ_{crit} (also shown in Fig. 1). In our new models, neutral gas is 'partitioned' into H I and H_2, and only H_2 is allowed to participate in SF. Therefore the value of Σ_{SFR} at a given $\Sigma_{\mathrm{H\,I+H_2}}$ has a 'second parameter' dependence. This second parameter is metallicity in the case of the GK (and KMT, not shown) recipes and disc mid-plane pressure (stellar surface density Σ_*, to first order) in the BR recipe. Fig. 6 shows the average metallicity of the cold gas in each galaxy, where each dot shows an annulus at a different radius, with a radial bin size of 500 pc. In the GK model, galaxies with higher metallicity have a higher Σ_{SFR} for a given $\Sigma_{\mathrm{H\,I+H_2}}$, as expected. However, we also see a similar dependence on metallicity in the BR model, although in this case it is not directly input into the model. The reason for this apparent dependence on metallicity is simply that Σ_* and gas phase metallicity are highly correlated. The curvature in Σ_{SFR}–$\Sigma_{\mathrm{H\,I+H_2}}$ at low gas surface densities in both models is in good agreement with observations of nearby spiral galaxies (Bigiel et al. 2008; Leroy et al. 2008).

3.3 Evolution of galaxy populations

3.3.1 Stellar mass functions and stellar fractions

In this subsection we examine the evolution of galaxy populations, which are directly comparable with observations. We focus on four main model variants: the fiducial versions of the GK, BR, and KS models, and the GK+Big1 model (see Table 2). The KS model is the SF recipe used in previously published Santa Cruz SAMs (S08; S12; Porter et al. e.g. 2014). The fiducial GK and BR models are the same as the models used in PST14 and Berry et al. (2014). For many of the quantities that we will show, the fiducial GK and BR models produce nearly indistinguishable results; in this case we note this and show only the GK model to avoid cluttering the plots. Selected results for other model variants are shown in Appendix B.

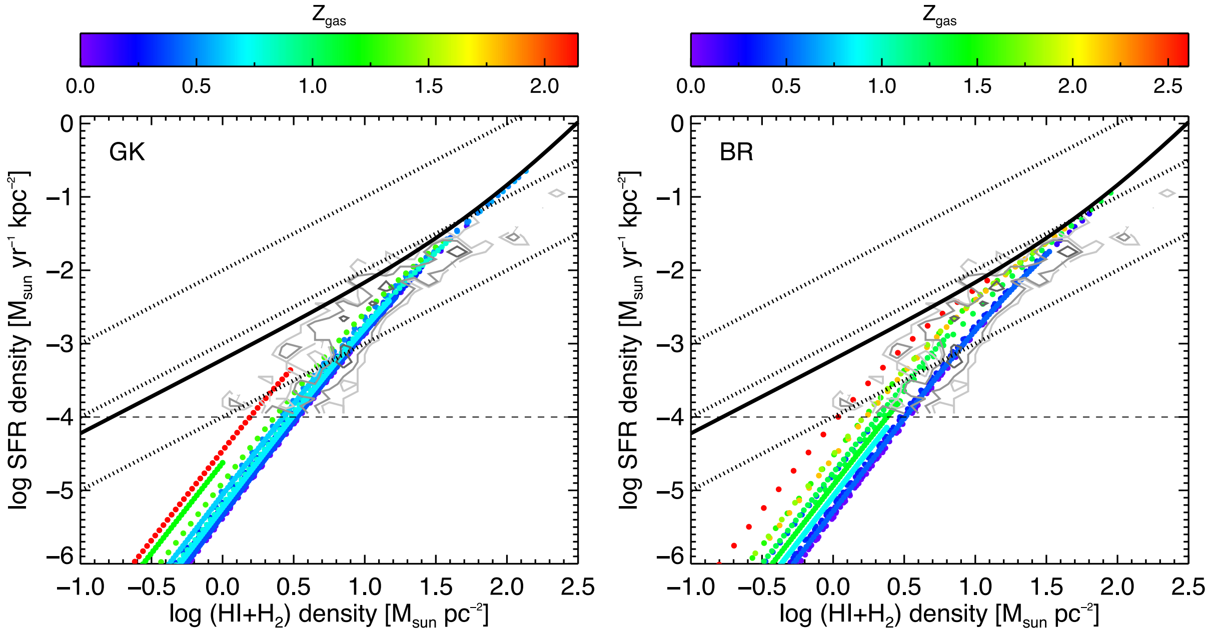

Figure 6. Relation between total cold neutral gas density (H I + H_2) and SFRD. The solid black line shows the input H_2-based SF recipe (Big2). Slanted dotted lines show an SFE of 1, 10, and 100 per cent per 10^8 yr. Grey contours show observational estimates from the 13 spiral galaxies from the THINGS+Heracles sample presented in Leroy et al. (2008). Coloured dots show a selection of 25 galaxies in our fiducial GK (left) and BR (right) models (at $z = 0$), with a stellar mass range chosen to match the THINGS sample. Each point shows the value in an annulus at different radius, with each annulus having a radial width $\Delta r = 500$ pc. The points are colour-coded with the average gas phase metallicity of the galaxy. Note that in both models, galaxies with lower metallicity gas have lower SFRD for a given total gas density, because a smaller fraction of the gas is predicted to be in the form of H_2. In the GK model, a direct dependence of H_2 fraction on metallicity is assumed. In the BR model, a dependence on disc mid-plane pressure is assumed, but this quantity turns out to be highly correlated with metallicity in our models.

In Fig. 7, we present predictions for the stellar mass function of galaxies from $z = 0$ to 6. We compare these predictions with a compilation of observations as described in the figure caption. We note that the currently available observational estimates of stellar masses and the corresponding stellar mass functions become rather uncertain above a redshift of $z \sim 1.5$, particularly at the faint end, where most surveys rely heavily on photometric redshifts, and surveyed areas are small, leading to a large uncertainty due to field-to-field variance (errors quoted in the literature typically do not include these uncertainties). These uncertainties become even larger at $z \gtrsim 4$ where the stellar masses are derived from the rest-UV part of the spectral energy distribution. More robust estimates should be obtainable from upcoming surveys. The fiducial BR model produces nearly indistinguishable results from those of the fiducial GK model, and is not shown here (it is shown for completeness in Fig. B1, in Appendix B). The similarity of the fiducial GK, BR, and KS model predictions is striking, particularly on the low-mass end where we might have expected the differences to be largest. The most noticeable differences are instead at high masses, particularly at redshifts $z \gtrsim 2$. Here, the KS model produces significantly lower number densities of massive galaxies at $z \gtrsim 1$, with the deviation growing with increasing redshift. It is important to note that we have not accounted for the expected errors in the observational estimates of the stellar masses when comparing to the models – this would tend to lead to an apparent increase in the number of massive galaxies due to Eddington bias (see e.g. Fontanot et al. 2009). However, even if this effect were included, the earlier formation of massive galaxies predicted by the GK and BR models is clearly in better accord with recent observations. In contrast, we see that the GK+Big1 model predicts a much lower number density of massive galaxies, with the difference between it and the fiducial models increasing with redshift. Even by $z \sim 1$, the discrepancy between the GK+Big1 model and observations is quite significant. Therefore this model variant appears to be disfavoured.

All three models suffer from the familiar excess of low-mass galaxies at $z \sim 0.5$–2 which, as we have already discussed, is a widespread problem in both SAMs and numerical hydrodynamic simulations (Fontanot et al. 2009; Weinmann et al. 2012; Somerville & Davé 2015; White et al. 2015). One of the important conclusions of this paper is that varying the SFE according to physically motivated recipes does not appear to be able to cure this problem within the current model framework. This is consistent with the conclusions of other recent studies (Fu et al. 2012; Henriques et al. 2013, 2015; White et al. 2015). A possibly related problem is that over this redshift range, the predicted gas fractions in low-mass galaxies in these same models may be too low (PST14; Somerville & Davé 2015; White et al. 2015; Popping et al. in preparation).[1]

Fig. 8 shows a related quantity, the stellar fraction (stellar mass divided by halo mass; $f_{\rm star} \equiv m_*/M_{\rm h}$) over the same redshift range. Our model predictions are now compared with constraints from (sub)-halo abundance matching from Behroozi et al. (2013). The conclusions are similar to the ones above, unsurprisingly since the $f_{\rm star}$ constraints are derived from observational estimates of stellar mass functions (though not exactly the same ones plotted in our Fig. 7). The median stellar fractions are nearly identical in the three 'fiducial' models (GK, BR, KS) in low-mass haloes ($\log M_{\rm h}/M_\odot \lesssim 11$) and are very similar in the GK and BR model over the whole range

[1] An important caveat, however, is that the gas mass estimates that lead to this conclusion are based on indirect methods. It remains to be seen whether this is conformed by direct observations of cold gas in these low-mass galaxies.

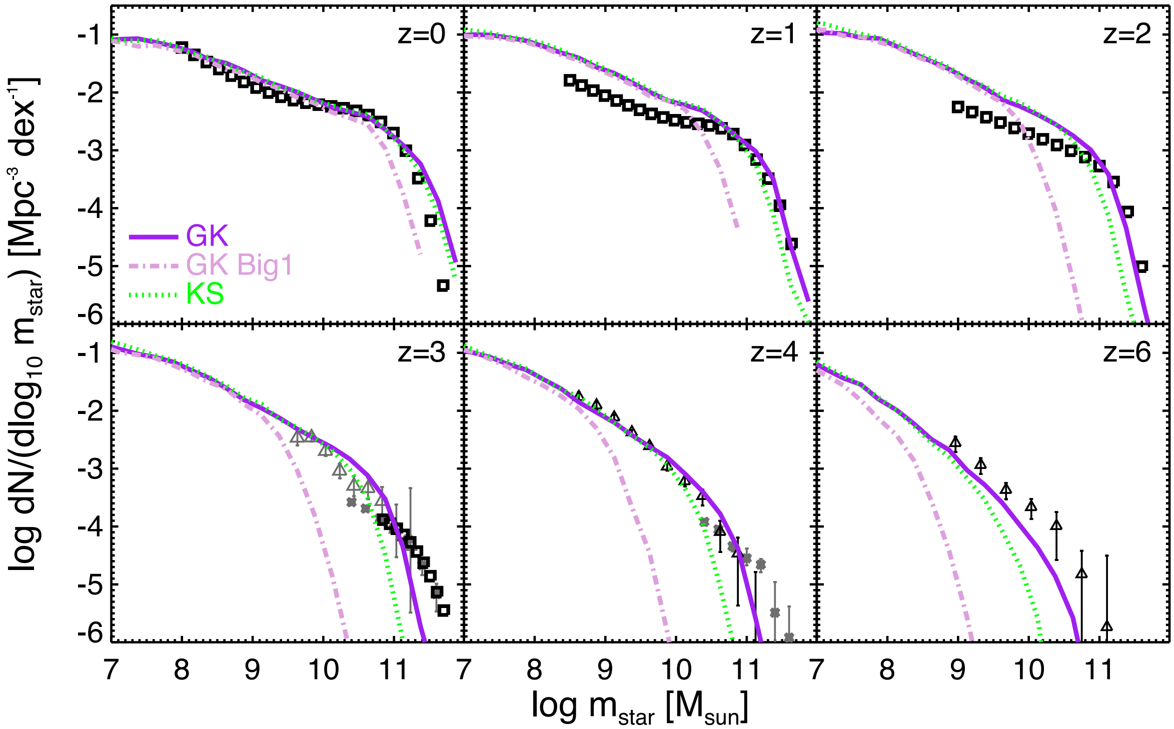

Figure 7. Stellar mass function evolution with redshift. Symbols show observational estimates as follows. In the $z = 0.1$, 1, 2, and 3 panels, black square symbols show a double-Schechter fit to a compilation of observational estimates. Observations included in the fit are: $z = 0.1$ – Baldry, Glazebrook & Driver (2008), Moustakas et al. (2013); $z = 1$ and $z = 2$ panels – Tomczak et al. (2014), Muzzin et al. (2013), $z = 3$ panels – Muzzin et al. (2013). The fits shown at $z = 1$, 2 and 3 are interpolated to these redshifts from adjacent redshift bins in the original published results. In the $z = 3$ panel, we also show estimates from Santini et al. (2012, triangles) and Caputi et al. (2011, crosses). In the $z = 4$ panel, we show estimates from Duncan et al. (2014, triangles) and Caputi et al. (2011, crosses). In the $z = 6$ panel, we show the estimates from Duncan et al. (2014, triangles). The purple solid line shows the results of the fiducial GK model, the lavender dot–dashed line shows the GK Big1 model, and the green dotted line shows the KS model.

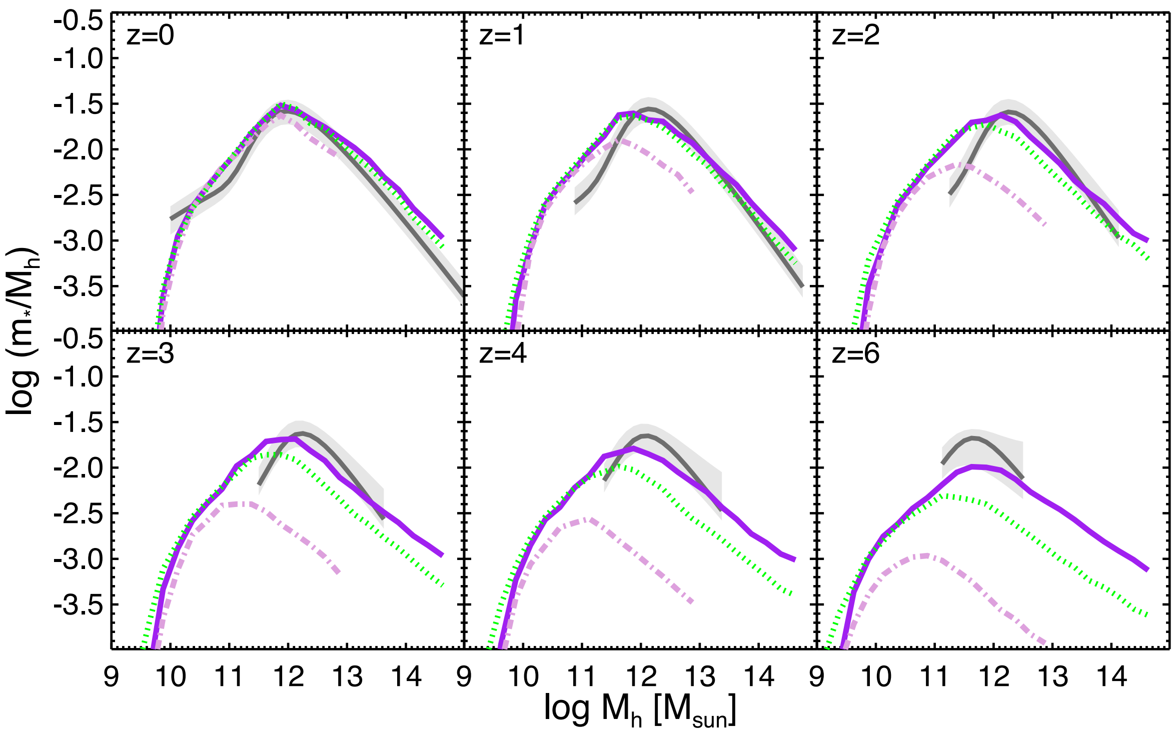

Figure 8. The stellar mass of central galaxies divided by the total mass of their dark matter halo, in redshift bins from $z = 0$–6. Dark grey lines and shaded areas show constraints from halo abundance matching from Behroozi, Wechsler & Conroy (2013). The purple solid line shows the results of the GK model, the orange dashed line shows the BR model, and the green dotted line shows the KS model. The lavender dot-dashed line shows the GK+Big1 model.

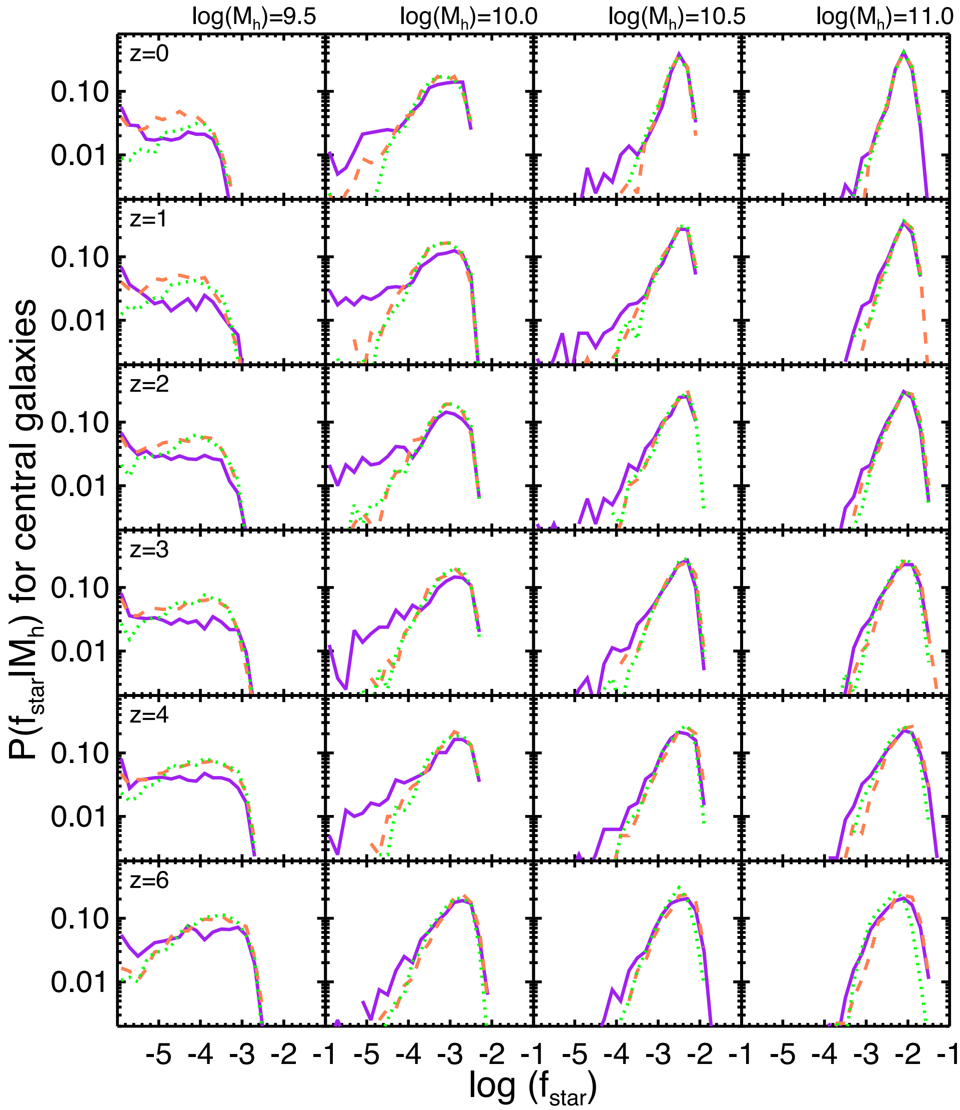

Figure 9. Distribution functions for the stellar fraction ($f_{\rm star} \equiv m_{\rm star}/M_{\rm h}$) of central galaxies in bins of halo mass ($\log M_{\rm h} = 9.25$–9.75, 9.75–10.25, 10.25–10.75, 10.75–11.25) and redshift as indicated on the panels. The purple solid line shows the results of the GK model, the orange dashed line shows the BR model, and the green dotted line shows the KS model.

of halo masses. The median value of $f_{\rm star}$ in the KS model is much lower in massive haloes than in the other two models, and the difference increases with redshift to about 0.4–0.5 dex at $z = 6$. The GK+Big1 model shows an even larger difference in the same sense, predicting much lower stellar fractions in high-mass haloes than the fiducial models, increasingly so at high redshift.

Although the median values of $f_{\rm star}$ are similar in the three fiducial models (GK, BR, KS), the distributions differ significantly for low-mass host haloes in these models. Figs 9 and 10 show the distribution of $f_{\rm star}$ in halo mass and redshift bins, for central and satellite galaxies, respectively. For all models and at all epochs, the distribution of $f_{\rm star}$ becomes broader and more skewed with decreasing halo mass. For massive haloes, the width of the distribution becomes slightly narrower with increasing time, while for low-mass haloes, a more noticeable tail towards lower values of $f_{\rm star}$ develops with time. In the two intermediate halo mass bins, this tail is more prominent in the GK models than in the other two models. The predicted broadening in $f_{\rm star}$ has potentially important implications for empirical halo-based models, which generally assume a narrow and fixed scatter in $f_{\rm star}(M_{\rm h})$. We show the results for these two types of galaxies separately because (a) central and satellite galaxies are treated differently in SAMs. For example, in our models, satellite galaxies are not allowed to accrete new gas from the IGM; (b) our predictions for differences between stellar fractions for satellites and centrals can provide useful input to empirical models such as halo occupation distribution and abundance matching models. Many such models do not distinguish between satellite and central galaxies.

3.3.2 SFR and gas depletion times

Fig. 11 shows the specific star formation rate (sSFR $\equiv \dot{m}_*/m_{\rm star}$) as a function of stellar mass over the redshift range $z = 0$–6. Our model predictions are compared with a compilation of observations as described in the figure caption. We have selected only 'star-forming' galaxies using the criterion sSFR $> 1/(3t_{\rm H}(z))$, where $t_{\rm H}(z)$ is the Hubble time at the galaxy's redshift. This has been shown to produce similar results to commonly used observational methods

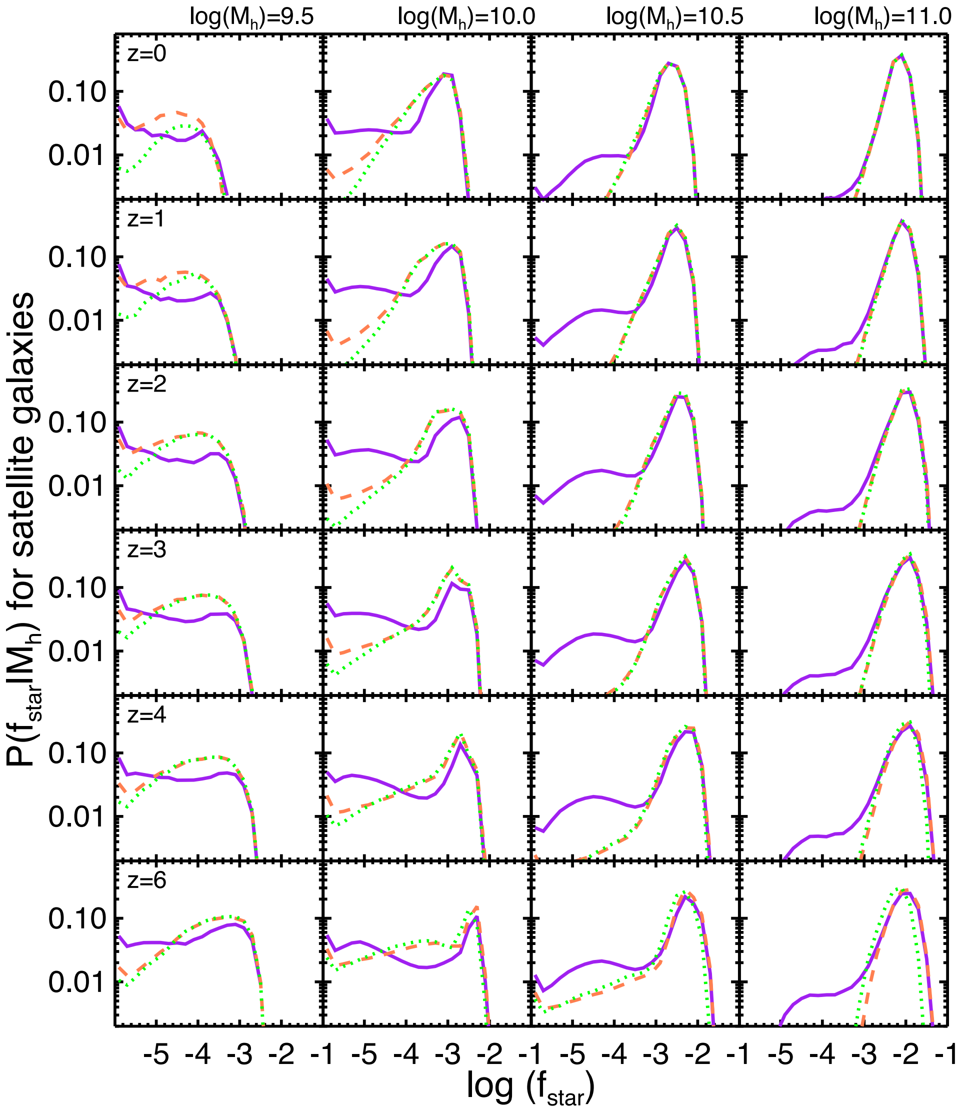

Figure 10. The same as Fig. 9, but for satellite galaxies (subhaloes). The halo mass here is the mass of the halo when it first becomes a subhalo.

for selecting star-forming galaxies (e.g. Lang et al. 2014). Our models agree well with the observed slope and normalization of the star forming main sequence (SFMS) at $z \sim 6$–4, but then the normalization of the model SFMS drops below the observationally estimated one between $z \sim 3$–0.5. At $z \sim 0$, the predicted SFMS has approximately the correct normalization for massive galaxies (here the precise value may be impacted by the details of the selection of 'star forming' versus quiescent galaxies), but the slope is much shallower than the observations suggest. This is another facet of the 'dwarf galaxy conundrum' discussed in White et al. (2015), and again is common to most cosmological models of galaxy formation (Somerville & Davé 2015). Our results show that this relation is extremely robust to changing the SF recipe in models. The median relations for the fiducial BR model are nearly indistinguishable from those for the fiducial GK model and are not shown. Fig. 12 shows the distribution of sSFR in stellar mass bins for the three fiducial models (GK, BR, KS). Even the sSFR distribution functions are very similar across these three models.

Fig. 13 shows the total gas depletion time $t_{\rm dep} \equiv (m_{\rm H\,I} + m_{\rm H_2})/\dot{m}_*$ in the fiducial GK and BR models, and in the GK+Big1 model. Here, we compute the depletion time using only the SFR due to the 'disc' mode of SF, i.e. not including SF due to merger-triggered bursts, but the plot looks very similar when the burst mode is included. Observations of nearby galaxies show that $t_{\rm dep}$ increases with decreasing stellar mass, i.e. the conversion of cold gas into stars is less efficient in low-mass galaxies (e.g. Leroy et al. 2008). A similar trend is indicated by the empirical estimates of total gas depletion time from Popping et al. (2015), also shown in Fig. 13 for comparison. The empirical estimates are based on an SFR–halo mass relation inferred from abundance matching, and an indirect estimate of the H I and H_2 masses from inverting the SFR density. These estimates rely on a number of assumptions (e.g. that disc cold gas radial profiles are well represented by exponentials), and on the observed relationship between size and stellar mass in disc-dominated galaxies. In Popping et al. (2015), the empirical predictions are shown only up to $z \sim 3$, because it is not known whether these assumptions and empirical relations hold at higher redshift. Here, we show the results of extrapolating the same method to $z \sim 6$, but these should be considered highly uncertain.

Our three fiducial models reproduce the same qualitative trends indicated by the observations and by the empirical predictions. First,

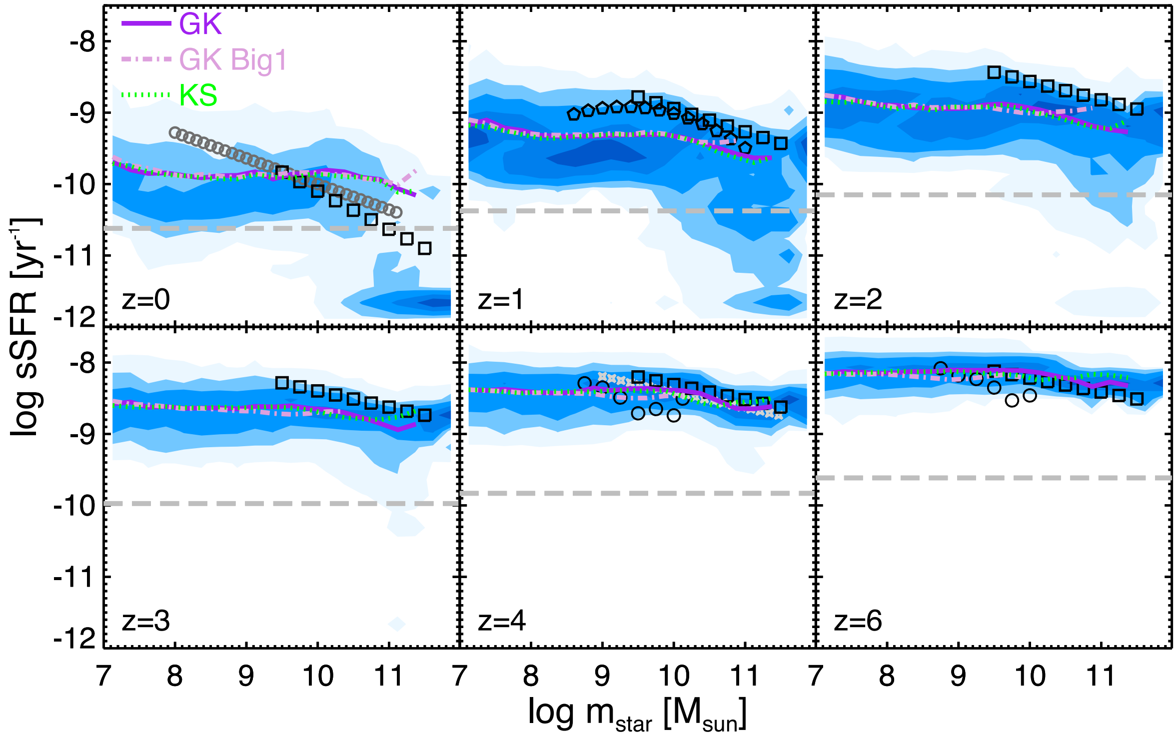

Figure 11. Mean specific star formation rate (sSFR $\equiv \dot{m}_*/m_*$), as a function of stellar mass for the fiducial GK (purple solid) and KS (green dotted) models, and for the GK+Big 1 model (lavender dot–dashed) in redshift bins from $z = 0$–6. The fiducial BR model is nearly indistinguishable from the GK and KS models, and is not shown. The blue contours show the conditional sSFR in the GK model. The horizontal grey line shows the sSFR corresponding to $1/(3t_{\rm H})$, where $t_{\rm H}$ is the Hubble time at that redshift. Only galaxies with sSFR $>1/(3t_{\rm H})$ are included in the mean. Symbols show a compilation of observations for star-forming galaxies as follows: $z = 0.1$ – Salim et al. (2007, open circles); $z = 1$ – Whitaker et al. (2014, pentagons, interpolated in redshift from the published results); $z = 4$ – Steinhardt et al. (2014, crosses); $z = 4$ and 6 – Salmon et al. (2015, circles); all panels – fit to data compilation from Speagle et al. (2014, squares).

depletion times are longer in lower mass galaxies. In detail, the physics that is responsible for this trend is different in the three fiducial models. Low-mass galaxies have lower gas and stellar surface density on average. In the KS model, gas below the critical surface density is not allowed to make stars, and low-mass galaxies tend to have a larger fraction of their gas below this threshold. Low-mass galaxies also tend to have lower gas phase metallicities, and in the GK model, this results in less efficient formation of H_2 and thus of stars. In the BR model, the lower SFE in low-mass galaxies is due to their lower stellar surface density, which leads to a lower disc mid-plane pressure, and again a lower H_2 fraction. Secondly, in all models, $t_{\rm dep}$ at a given stellar mass was lower in the past and increases with cosmic time. The GK and BR models show more pronounced evolution and shorter depletion times (higher SFE) at high redshift, particularly in massive galaxies. This is due to the steeper dependence of the SFR density on gas density adopted in our GK and BR models (slope $N_{\rm SF} = 2$ instead of 1.4). High-redshift galaxies contain higher surface density gas overall, and so the results are more sensitive to the slope of the SF relation at high gas densities. This result explains the less efficient formation of stars in massive galaxies at high redshift in the KS model relative to the GK and BR models, seen in Figs 7 and 8. Note that already by $z \sim 3$, and increasingly so at higher redshift, the predictions using the Big1 recipe for SF are inconsistent with the empirical constraints. This suggests that the assumption of a constant H_2 depletion time in galactic discs (which is inherent in the Big1 recipe) may not be universally applicable.

Fig. 14 shows the H_2 depletion time ($t_{\rm dep,H_2} \equiv m_{\rm H_2}/\dot{m}_*$) in the GK, BR, and GK+Big1 models (the KS model is not shown, as we do not track H_2 self-consistently in this model). This figure, in combination with results shown in PST14, shows that the trends seen in our models in Fig. 13 are due to a combination of two factors: (a) at a given redshift, more massive galaxies have larger fractions of their cold gas in the form of H_2, and at fixed mass, higher redshift galaxies have higher H_2 fractions; (b) the H_2 depletion time is also shorter in more massive galaxies and at high redshift.

3.3.3 Mass–metallicity relations

Figs 15 and 16 show the mass–metallicity relation (MZR) for stellar and cold gas phase metallicities, respectively. Recall that the chemical yield parameter in our models has been adjusted to approximately reproduce the normalization of the stellar MZR measured by Gallazzi et al. (2005). Our models naturally predict a slope for the stellar MZR that is in fairly good agreement with observations (Woo, Courteau & Dekel 2008; Kirby et al. 2013) down to very low stellar masses ($m_{\rm star} \sim 10^7\,{\rm M}_\odot$). The three fiducial models make very similar predictions for the stellar phase MZR, except that massive galaxies become enriched much earlier in the fiducial GK and BR models than in the KS model (the predictions of the fiducial BR model are nearly identical to those of the fiducial GK model and are not shown). This is owing to the steeper slope of our Big2 SF recipe at high gas densities (see discussion in Section 3.1, especially Fig. 2, and above). This is further illustrated by comparing the

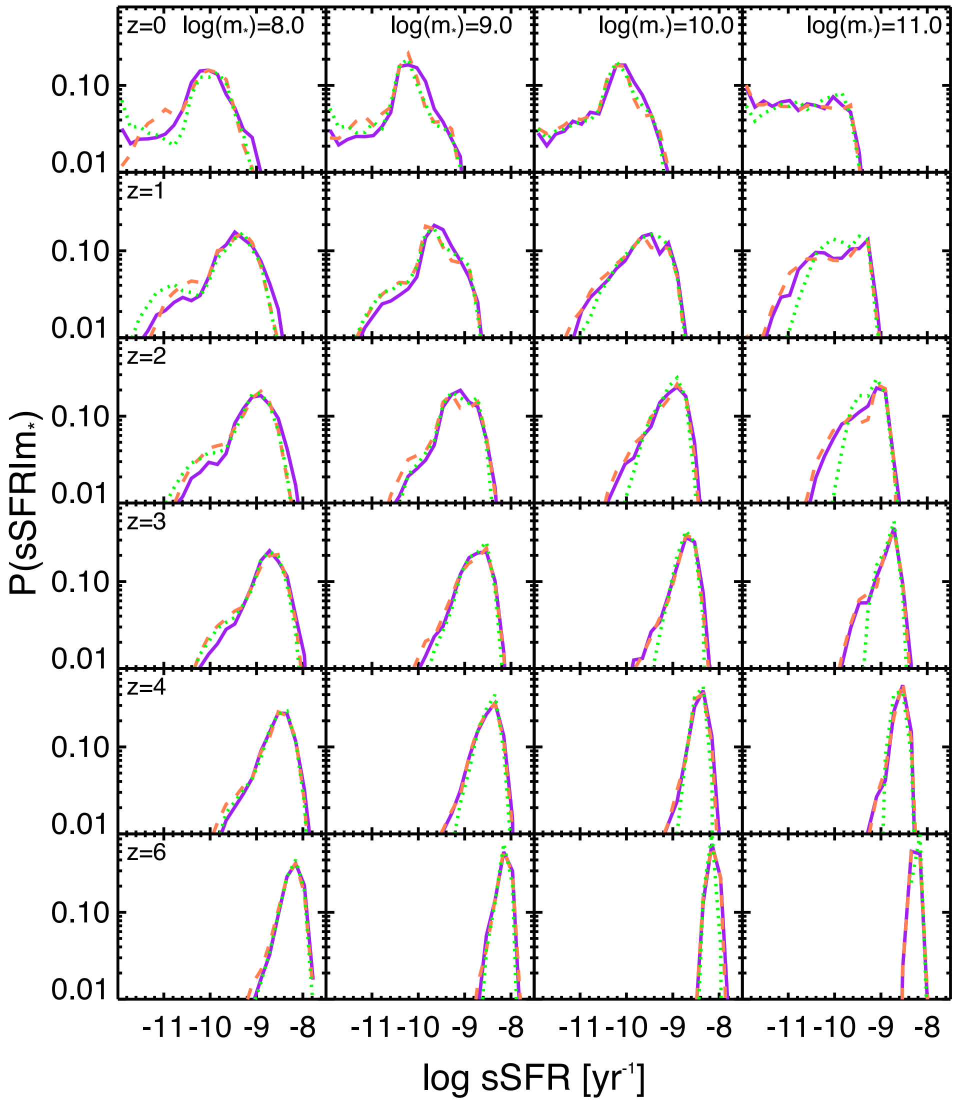

Figure 12. The conditional probability distributions of sSFR in stellar mass bins at different redshifts, for our three fiducial models (purple solid: GK; orange dashed: BR; green dotted: KS).

predictions for the GK+Big1 model, in which enrichment is delayed even further. It is also interesting to note that our models predict a much smaller dispersion in the stellar MZR than the observational dispersion estimated by Gallazzi et al. (2005).

In Fig. 16, the model results shown are for central, star-forming galaxies selected using the same criteria described above. The sSFR cut is applied because the observational estimates of gas phase metallicity are based on emission line diagnostics from H II regions, which are only detectable in star-forming galaxies. We omit satellite galaxies from the comparison because satellite galaxies are known to be overquenched and overly gas poor in our current models. As a result, their MZR is offset from that of the central galaxies. A compilation of observational estimates for the gas phase MZR is shown. We have converted the observed values of $12 + \log(\mathrm{O/H})$ to $Z/\mathrm{Z}_{\odot}$ assuming $12 + \log(\mathrm{O/H}) = 8.76$ for the Sun (Caffau et al. 2011), and $\mathrm{Z_g/Z_{\odot}} = (\mathrm{O/H})/(\mathrm{O/H})_{\odot}$. Note that the observed MZR is uncertain by a factor of 2–3 as different calibration methods produce different zero-points and, to some extent, different slopes (Kewley & Ellison 2008). The results are again quite similar, with the KS model again producing later enrichment of massive galaxies, so the MZR is shallower at $z \sim 3$–6, and the GK+Big1 model producing even more delayed enrichment. All three models produce a $z = 0$ cold gas phase MZR that is, taken at face value, considerably steeper than the observed gas phase MZR. Moreover, the predicted *evolution* of gas phase metallicity in our models is quite different from that implied by current observations. The models predict that the gas phase metallicity for galaxies of fixed stellar mass declines slightly with decreasing redshift, while observations indicate an increase of almost a factor of 2 between $z \sim 2.2$ and 0. This discrepancy was shown previously by White et al. (2015) for our KS models; we see here that the results are qualitatively similar for our new fiducial GK and BR models. We discuss possible reasons that our models reproduce the stellar MZR fairly well but seem to fail to reproduce the observed gas phase MZR in Section 4.2.

3.3.4 *Evolution of global quantities*

In Fig. 17, we show our model predictions for the evolution of the global SFRD, global stellar mass density, and average metallicity

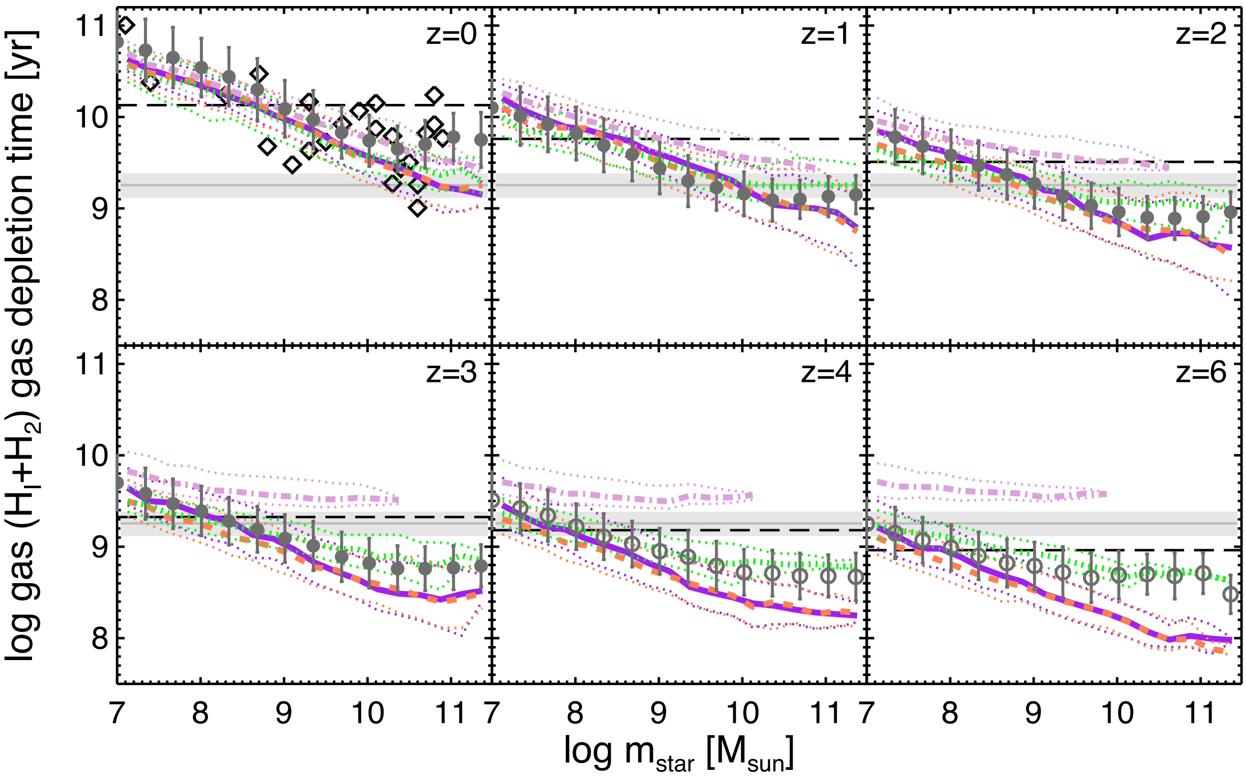

Figure 13. Total gas depletion time, defined as the total cold neutral gas mass (H I + H_2) divided by the SFR. The purple solid lines show the predictions of the fiducial GK model, the orange dashed lines show the BR model, and the green dotted lines show the KS model. The lavender dot–dashed lines show the GK+Big1 model. The 16th, 50th, and 84th percentiles are shown for central star-forming galaxies in the models. The horizontal black dashed line shows the age of the Universe at that redshift. The open diamonds in the $z = 0$ panel show observational estimates for galaxies in the THINGS+Heracles sample from Leroy et al. (2008). The horizontal grey line shows the average *molecular* gas depletion time estimated by Leroy et al. (2013) for nearby galaxies; the shaded grey area indicates the uncertainty in the measurement due to the uncertain conversion factor between CO and H_2. The grey circles show the estimates obtained via the empirical method of Popping, Behroozi & Peeples (2015). These are plotted with open symbols at $z > 3$ to indicate that the estimates are quite speculative in this redshift regime (see text for more details). All three models reproduce the observed trend of decreasing depletion time with increasing stellar mass, but different underlying physics are responsible for the trends in the different models.

of cold gas and stars over cosmic time. The figure shows that the three fiducial models predict almost identical global SFRD at low redshift, while at $z \gtrsim 2$, the BR model produces the highest SFRD, and the KS model the lowest, with the GK model in between. Although the GK+Big1 model also predicts a similar SFRD at $z \sim 0$, it produces significantly later SF, assembly of stellar mass, and chemical enrichment, and the predicted global SFR density is much lower than the observational constraints at $z \gtrsim 1$. We also see that our fiducial model predictions are in reasonably good agreement with observations at $z \lesssim 2$ and $z \gtrsim 6$, though this is in part a fortuitous cancellation – the models overproduce galaxies with low SFR but underproduce ones with high SFR. Moreover, the observational estimates of the SFRD from Madau & Dickinson (2014) are integrated only down to 0.03 L_*, while our theoretical predictions are integrated over all galaxies. The Khaire & Srianand (2015) SFRD estimates are integrated down to zero luminosity and so should be more comparable to our model predictions. This same behaviour is echoed in the build-up of the global stellar mass density. The largest difference in the three fiducial models appears in the evolution of the stellar and cold gas phase metallicity. The models differ in the normalization and evolution of the mean metallicity for gas and stars. In addition, in the KS model, the mean metallicity of cold gas and stars is very similar, while there is a much larger difference between the stellar and gas phase metallicities of gas and stars in the BR model and GK models. As one can see in Figs 15 and 16, at stellar masses above $m_{\rm star} \sim 10^7$ M$_\odot$, the models make similar predictions for the relative metallicities of gas and stars. The differences seen in Fig. 17 are entirely due to very low mass haloes.

4 DISCUSSION

4.1 Interpreting our results: the equilibrium model

One of the main conclusions of our work is that modifying the recipes for how cold gas is converted into stars has very little effect on the properties of low-mass galaxies ($m_* \lesssim M_{\rm char}$, where $M_{\rm char}$ is the 'knee' in the stellar mass function). Instead, modifying the SF recipe mainly changes the ratio of cold gas to stars in galaxies. Similar conclusions were reached in the study by White et al. (2015), in which more extreme (though in some cases less physically motivated) modifications to the SF recipe in similar models were made. This is because in our models, SF in low-mass galaxies is strongly self-regulated: if SF is made less efficient, less gas is ejected by stellar winds, leading to more efficient SF, and vice versa.

A number of recent works have pointed out this property of self-regulation, which is a rather generic feature of modern galaxy formation models, arising from the broadly adopted hypothesis of

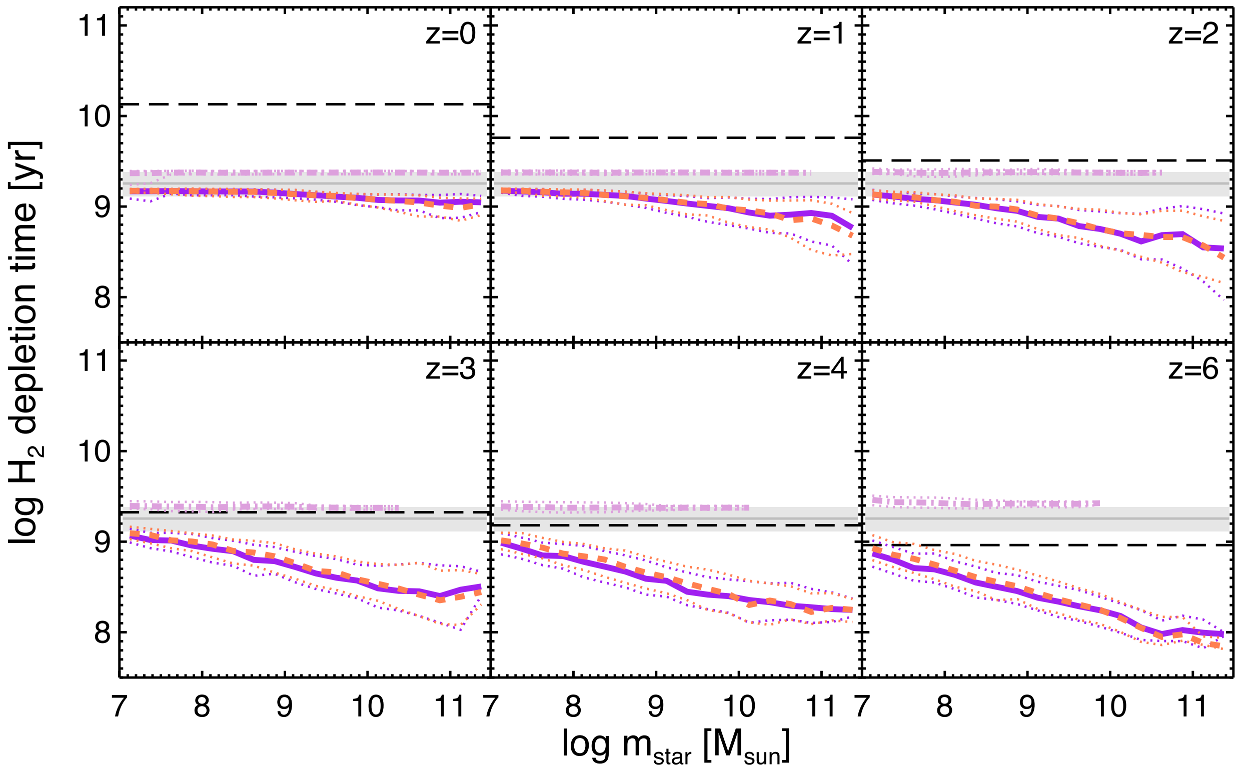

Figure 14. H_2 depletion time, defined as the H_2 mass divided by the SFR. Models shown (coloured lines) are as in Fig. 13. The horizontal black dashed line shows the age of the Universe at that redshift. The horizontal grey line shows the average *molecular* gas depletion time estimated by Leroy et al. (2013) for nearby galaxies; the shaded grey area indicates the uncertainty in the measurement due to the uncertain conversion factor between CO and H_2. Our models predict that at high redshift, molecular gas depletion times were significantly shorter, and there is a much stronger trend between galaxy stellar mass and H_2 depletion time.

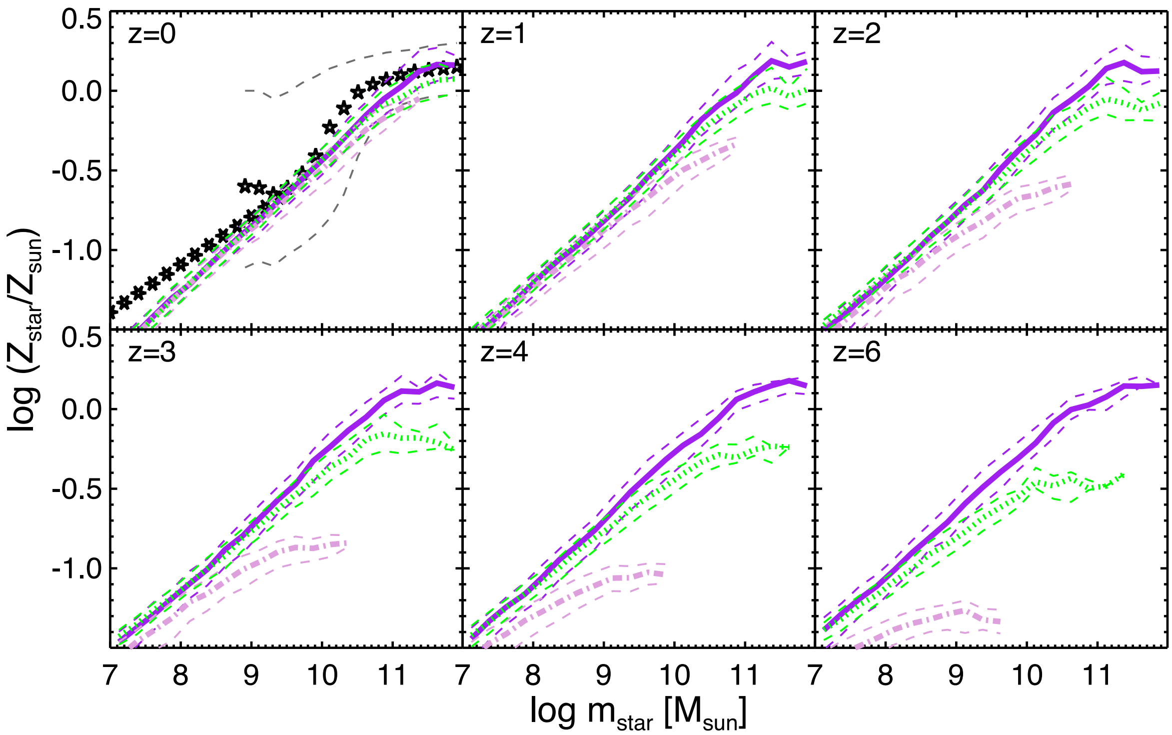

Figure 15. Stellar mass versus stellar metallicity. Five-pointed star symbols and dashed lines show observational estimates from Gallazzi et al. (2005), and six-pointed stars show the fit to the observed MZR estimated from Local Group dwarf galaxies by Kirby et al. (2013). The purple lines show the 16th, 50th, and 84th percentiles for our fiducial GK model, the lavender lines shows the GK+Big1 model, and the green lines show the KS model. The results of the fiducial BR model are nearly indistinguishable from the GK model and are not shown.

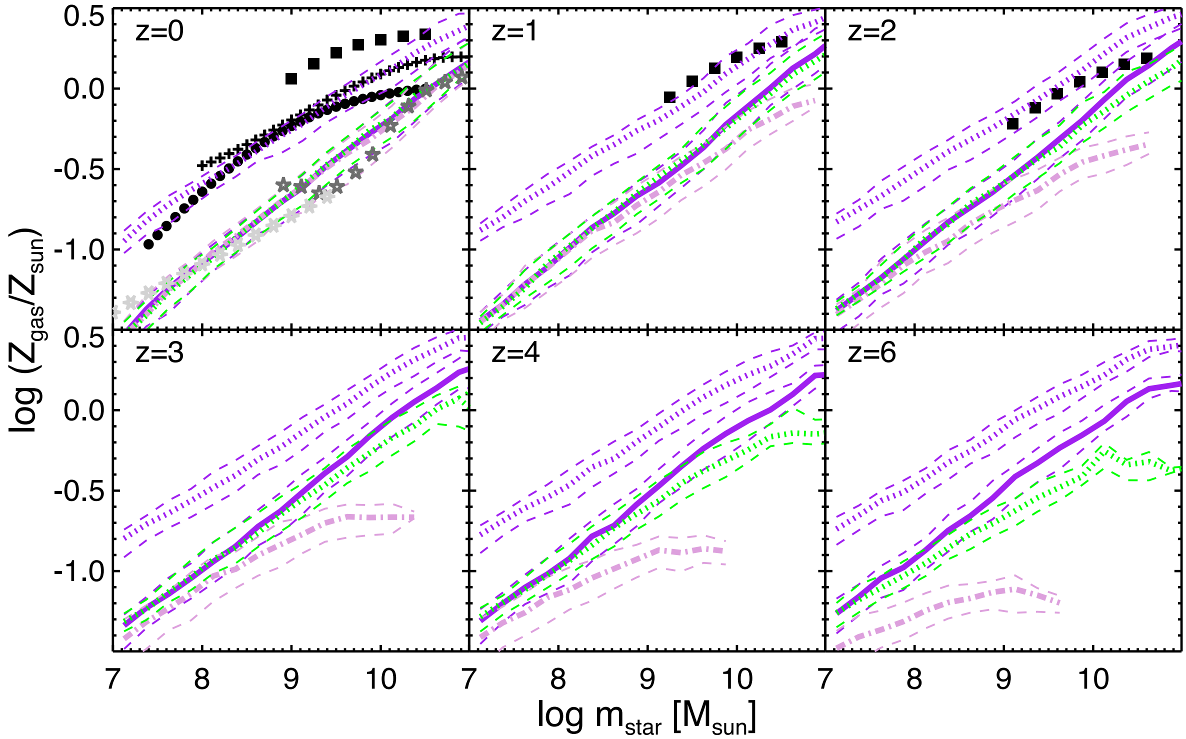

Figure 16. Stellar mass versus cold gas phase metallicity. Black symbols show observational estimates of the gas phase metallicity: $z = 0.1$ – Peeples et al. (2014, pluses); Andrews & Martini (2013, filled circles). In all panels, the filled squares show the compilation of observational estimates from Zahid et al. (2013). For comparison, we also show the observational estimates of stellar metallicity, as in Fig. 15, with grey star symbols. Coloured lines show model predictions, as in Fig. 15. The dotted purple line shows the fiducial GK model with an approximate correction for varying [α/Fe] (see text). Taken at face value, our predicted gas phase MZR appears to be much steeper than the observational estimates. However, properly accounting for varying abundance ratios of α versus Fe elements may at least partially remove this tension (see text).

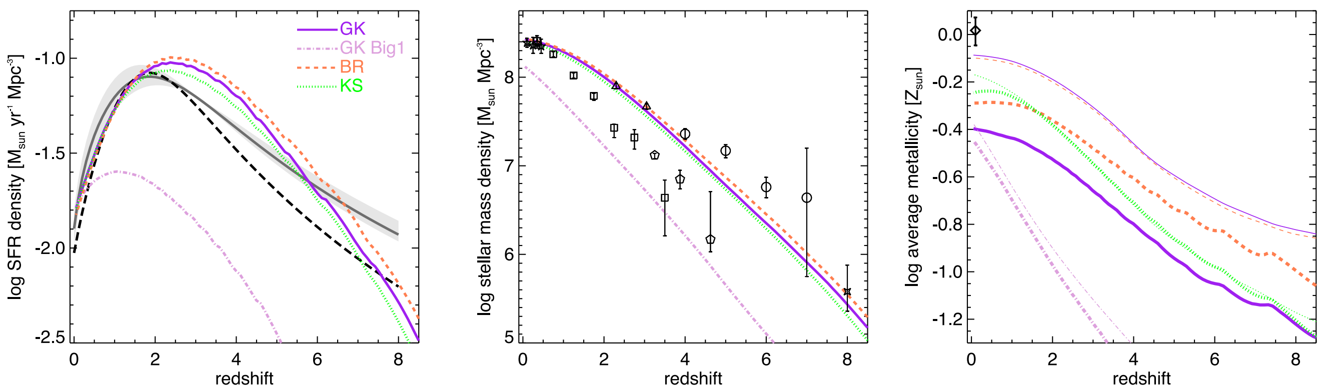

Figure 17. Global history of SF, stellar mass assembly, and metallicity as a function of redshift. The solid purple line shows the results of our fiducial GK model, the dashed orange line shows the BR model, the dotted green line shows the KS model, and the dot–dashed lavender line shows the GK+Big1 model. Left: global SFRD; the long dashed black line shows the fit to the compilation of observational estimates from Madau & Dickinson (2014), and the solid grey line and shaded area shows the observational estimates of the median, low, and high SFRD from Khaire & Srianand (2015, their LMC2 dust model). Middle panel: global stellar mass density; symbols show selected observational estimates taken from table 2 of Madau & Dickinson (2014). Right: mean mass-weighted metallicity; thick lines show the metallicity of the cold gas component and thin lines show the stellar component. The diamond symbol at $z \sim 0$ is the mean stellar metallicity derived by Gallazzi et al. (2005).

efficient stellar-driven feedback (Schaye et al. 2010; Haas et al. 2013; Somerville & Davé 2015; White et al. 2015). A useful analytic framework for understanding the behaviour of the fairly complex intertwined suite of physical processes at play in state-of-the-art galaxy formation simulations is the 'equilibrium model', sometimes called the 'bathtub model' (e.g. Davé, Finlator & Oppenheimer 2012; Dekel et al. 2013; Dekel & Mandelker 2014). The basic assumption in this model is that due to self-regulation, on some time-scale $t_{\rm eq}$, galaxies establish an equilibrium state in which the rate of change of their cold gas reservoir is small, i.e. $\dot{m}_{\rm cold} \simeq 0$. Once equilibrium is established, the SFR is balanced by global inflows and outflows, $\dot{m}_* = \dot{m}_{\rm in}/(1 + \eta)$, where $\dot{m}_{\rm in}$ is the rate at which gas flows into the galaxy due to cosmological accretion and $\eta \equiv \dot{m}_{\rm out}/\dot{m}_*$ is the mass loading factor of a large-scale stellar driven outflow. The time for a galaxy to come into equilibrium (or to re-establish equilibrium after a disruption) is

$$t_{\rm eq} = \frac{\tau_{\rm SF}}{1+\eta}, \tag{9}$$

where $\tau_{\rm SF} \equiv m_{\rm cold}/\dot{m}_*$ is the SF (or gas depletion) time-scale (Davé et al. 2012, hereafter DFO12).

In simulations, η is a fairly strong inverse function of galaxy mass, while $\tau_{\rm SF}$ is a weaker function of galaxy mass (DFO12). Therefore low-mass galaxies come into equilibrium earlier. This helps us to understand why changing our SF recipe had less impact on low-mass versus massive galaxies. In addition, it explains why high-mass galaxies were affected only at high redshifts – at $z \gtrsim 2$, these galaxies have not yet come into equilibrium. Furthermore, in models that use the Big1 recipe for converting H_2 into stars, the gas depletion times at high redshift are considerably longer than in models adopting the Big2 recipe, especially for massive galaxies. Therefore these massive galaxies come into equilibrium even later in the Big1 model. This work makes the interesting prediction that observations of massive galaxies at very high redshift ($z \gtrsim 2$) will place strong constraints on the physics of SF in cold dense gas, while constraints on the low-mass end of the galaxy stellar mass function at high-z will mainly constrain the physics of outflows.[2]

4.2 Mass–metallicity relations for gas and stars

It is puzzling that our models reproduce the stellar MZR fairly well but then predict a much steeper gas phase MZR than observations appear to indicate. This is seen in other models as well – models that invoke a weaker dependence of mass outflow rate on galaxy circular velocity, and normalize their yield parameter to the observed *gas phase* MZR, produce better agreement with the observed gas phase MZR but then fail to reproduce the stellar MZR (Lu et al. 2014). We can see by comparing Figs 15 and 16 that our models predict that the stellar metallicity in galaxies is about a factor of 1.5–1.7 lower than the gas phase metallicity, with weak trends on stellar mass and redshift. We can see from the figures in Section 3.1 that the gas phase metallicity tends to increase rapidly and monotonically with time in our models. The stellar metallicity is effectively a mass-weighted average over the chemical enrichment history of the galaxy, so it makes sense that the stellar metallicities are slightly, but not enormously, lower than the gas phase metallicities at any given time. Taken at face value, the observational results – which imply that $Z_{\rm gas}/Z_{\rm star}$ is as high as a factor of ∼10 or more, and is a fairly strong function of stellar mass – may be difficult to reproduce in cosmological models without invoking accretion of highly metal pre-enriched gas.

An alternative explanation is that the normalization, and possibly the slope, of the gas and stellar phase MZRs are not accurately calibrated to the same system. Indeed, some gas phase metallicity indicators do yield a gas MZR normalization and slope that is more consistent with the stellar MZR (Kewley & Ellison 2008). Another issue is that different observational probes are sensitive to different chemical elements. The stellar metallicities derived by Gallazzi et al. (2005) are sensitive to a combination of Fe and Mg, and the stellar MZR derived by Kirby et al. (2013) measures [Fe/H]. Gallazzi et al. (2005) quote their results in terms of $Z_{\rm star}/Z_\odot$ and claim that their results are independent of α/Fe (Gallazzi, private communication). The observational gas phase MZRs are primarily sensitive to α elements and are quoted in terms of oxygen abundance (12 + log (O/H)). We have assumed that $Z_{\rm gas}/Z_\odot = ({\rm O/H})/({\rm O/H})_\odot$ and $Z_{\rm star}/Z_\odot = ({\rm Fe/H})/({\rm Fe/H})_\odot$ for the Kirby et al. (2013) observations. This is equivalent to assuming that all the stars in our model galaxies have $(\alpha/{\rm Fe}) = (\alpha/{\rm Fe})_\odot$. However, $(\alpha/{\rm Fe})$ is known to differ significantly from the Solar value in stars in our own Galaxy (e.g. Stoll et al. 2013), in nearby dwarf galaxies (Tolstoy et al. 2009, and references therein), and in giant ellipticals (Trager et al. 2000; Thomas et al. 2005). If we assume that the metallicity tracked in our models is actually Fe, and apply the empirical relation presented by Stoll et al. (2013)[3] to our model galaxies to 'convert' to [O/H], we find much better agreement between our predicted gas phase MZR and at least some calibrations of the observed MZR (see Fig. 16). Note that, conceptually following Muñoz & Peeples (2015), we are effectively assuming that a relationship between [Fe/H] and [O/H] derived for individual stars in the Milky Way holds for the *average* stellar and gas metallicities in galaxies with a variety of SF histories – an assumption that may well not be valid. However, it suggests that properly accounting for non-Solar α/Fe and its possible trends with other galaxy properties (such as stellar mass and metallicity) may at least partially relax the tension between the stellar and gas phase MZR seen in our models and others.

In the models presented here, we track the total metallicity assuming a constant yield, and we also adopt the instantaneous recycling approximation. Metals are produced in direct proportion to the formation of stars, so enrichment in our models probably most closely traces α elements, but we normalized our yield parameter to observations that are also sensitive to Fe (see above). This simple version of single-element chemical enrichment with the instantaneous recycling approximation is the standard approach adopted in SAMs. However, physical processes in the models are actually dependent on different elements in potentially significant ways. For example, the (also widely adopted in SAMs) cooling tables of Sutherland & Dopita (1993) used to model the cooling rate of hot halo gas are parametrized via [Fe/H] and adopt assumed [Fe/H]-dependent abundance ratios. However, H_2 formation is more closely tied to α elements such as oxygen and carbon, which are primary coolants in the ISM (e.g. Glover & Clark 2014). One conclusion of the work presented here is that if we wish to include more realistic physics in our models, it is important to track multiple chemical elements and their production via different channels (stars of various masses, Type II SNe and prompt and delayed Type Ia SNe) and on different time-scales. Several groups have developed SAMs that include more detailed chemical evolution models (Arrigoni et al. 2010; Yates et al. 2013). We have integrated the more sophisticated chemical evolution models presented in Arrigoni et al. (2010) within our new SAMs, including the new metallicity-dependent H_2 formation and H_2-based SF recipes as described here, and plan to investigate the predicted metallicities of cold gas and stars and their evolution in these models in a future work (Peeples et al. in preparation).

[2] All of this discussion implicitly assumes that 'preventative' feedback – physical processes that could prevent gas from accreting into galaxies or becoming available for SF – is sub-dominant. At very low masses, preventative feedback due to photoionization squelching likely becomes important. Preventative feedback due to AGN heating and winds, and virial shock heating, is probably dominant in massive galaxies at late times ($z \lesssim 2$). See DFO12 and Somerville & Davé (2015) for a more complete discussion.

[3] $[{\rm Fe/H}] = -0.34 + 1.25[{\rm O/H}]$.

4.3 Caveats and limitations of our models

Understanding how the 'small-scale' processes of SF and stellar feedback interact with cosmological scale processes such as galactic scale inflows and outflows to shape the observable properties of galaxies is currently one of the major unsolved problems in the study of galaxy formation and evolution. The models presented here neglect a large number of processes that are thought to be important in influencing how efficiently gas can form molecules, and in turn how stars form within molecular gas. For example, we do not consider the possible impact of the local shear field, and non-axisymmetric perturbations such as spiral arms and bars. Nor do we attempt to model the 'local' effects of stellar feedback (through stellar winds, SNe, and H II regions) on SF. We instead assume that the efficiency of converting *molecular gas* into stars is roughly constant, as suggested by observations of nearby spiral galaxies (Bigiel et al. 2008, 2011). As pointed out by Krumholz, Dekel & McKee (2012) and many others, the efficiency of forming stars *within* GMCs is surprisingly low, only about 1 per cent per free fall time. Our picture is that local feedback processes are responsible for setting that efficiency, and that when we smooth over several 100 pc regions of the ISM, it averages out to a nearly universal value.

However, there have been some recent studies that suggest that the galaxy-averaged value of this molecular SFE (often expressed as a depletion time, $t_{\rm dep,H_2} \equiv m_{H_2}/\dot{m}_*$) may vary significantly from galaxy to galaxy, and may have a strong dependence on global galaxy properties. Saintonge et al. (2011) find that in the COLD GASS sample, $t_{\rm dep,H_2}$ is weakly correlated with the galaxy stellar mass and stellar surface density, and rather strongly (anti-) correlated with sSFR. However, Saintonge et al. (2011) adopted a constant (Galactic) value for the conversion factor between CO and H_2 ($\alpha_{\rm CO}$). Leroy et al. (2013) found similar correlations in their sample of 30 disc galaxies from the HERACLES survey, but found that most of the correlation disappeared when they applied a theoretically motivated dependence of $\alpha_{\rm CO}$ on the dust-to-gas ratio. They found smaller residual variations in $t_{\rm dep,H_2}$, mainly associated with nuclear gas concentrations. This suggests that there may be different values of $t_{\rm dep,H_2}$ in undisturbed disc galaxies and in galaxies experiencing mergers and interactions (see also Daddi et al. 2010; Genzel et al. 2010), possibly due to a superlinear dependence of SFRD on H_2 density, as assumed in our fiducial models. We do include an enhancement of SFE in mergers in our models, but the treatment is based on hydrodynamic simulations of binary mergers with a rather outdated treatment of sub-grid physics, so this is clearly an area that should be explored with state-of-the-art high-resolution hydrodynamic simulations.

Another limitation of our approach is that although we compute the H_2 fraction and $\Sigma_{\rm SFR}$ in radial annuli in each disc, and integrate over the disc to obtain the global properties, we do not store the information on the stellar, gas, and metal content of each annulus over different timesteps. Therefore we assume that both the gaseous and stellar discs have radial exponential profiles, and adopt a simple fixed factor relating the size of the gaseous and stellar disc. While this may be a reasonable approximation on average, it may miss important trends. We also use the global values of the gas phase metallicity and SFR (which we use as a proxy for the UV radiation field) in the GK recipe, instead of the local values of these quantities in annuli. In future work, we plan to construct more detailed models of discs in which all of these quantities are tracked as a function of radius, along the lines of work by Fu et al. (2010) and Dutton et al. (2010).

4.4 Comparison with previous work

Several other groups have carried out studies similar to this one. Lagos et al. (2011b) considered two models without gas partitioning and two models with gas partitioning. The first unpartitioned model adopts the original SF relations implemented in the Baugh et al. (2005) and Bower et al. (2006) GALFORM models, in which the SFR was assumed to be proportional to the total cold gas mass divided by a time-scale τ_*. In the Baugh et al. (2005) models, τ_* was scaled with the galaxy circular velocity to a power, and in the Bower et al. (2006) models, also with the galaxy dynamical time. The second unpartitioned model used a KS SF recipe similar to the one we have adopted here. The two models with gas partitioning adopted either the pressure-based BR recipe or the metallicity-based KMT recipe. In their BR model, they adopted SF relations similar to our Big1 and Big2 recipes, which they call BR and BR.nonlin. In the KMT model, they used the SF relation given by KMT. They did not attempt to separate the effects of the different gas partitioning recipes versus H_2-based SF recipes. We focus on the results of the Bower et al. (2006) variant of the GALFORM models, which are more similar to our models than the Baugh et al. (2005) variant. Lagos et al. (2011b, hereafter L11a) do not show stellar mass function predictions, but find that the rest-frame *K*-band luminosity function at $z = 0$, 1, and 2 shows little change between the six different combinations of SF and gas partitioning recipes they explored. This is mostly consistent with our results, but a few points are worth noting. L11a show only the BR gas partitioning recipe in combination with a Big1-like SF recipe, and do not show it in combination with a metallicity-dependent gas partitioning recipe. We find that our BR+Big1 model looks very similar to our fiducial GK model (see Fig. 5), which seems consistent with the lack of variation seen by L11a. This suggests that the delayed enrichment combined with a metallicity-dependent gas partitioning recipe in the GK+Big1 model is largely responsible for the large differences between that model and our other models. It is also worth noting that our Big2 SF recipe steepens to a slope of 2 at the highest H_2 surface densities, while the L11a BR.nonlin model has a maximum slope of 1.4, similar to the KS recipe. This may also lead to larger differences between our Big1 and Big2 models relative to the BR and BR.nonlin models of L11a. However, our implementation of the KMT recipe should be very similar to the KMT recipe of L11a, and yet they seem to find a smaller change in the bright end of the luminosity function (LF) between KMT and their other models than we see in the high-mass end of the stellar mass function for our fiducial GK and BR models compared with our KMT model (see our Fig. B1). This could either be due to a different evolution in the *K*-band LF arising from evolving mass-to-light ratios, or to differences in other physical processes in the models.

L11a emphasize differences in the SFR distributions as a function of stellar mass, in particular the prominence of a passive population in models with different SF recipes. We find very small differences in the SFR distributions between our different models, and note that the prominence and location of a passive population will certainly also be very sensitive to the treatment of AGN feedback. It is also interesting to note that Lagos et al. (2011a) end up favouring their BR models and strongly disfavour the KMT recipe because they find that it does not reproduce the observed H I and H_2 gas scaling relations. In contrast, we showed in PST14 that both our BR and GK models do about equally well at reproducing available observations of cold neutral gas in galaxies. Lagos et al. (2012) show predictions for the relationship between *B*-band luminosity and cold gas phase metallicity at $z = 0$ in their fiducial model, which adopts a

BR-like pressure-based recipe for gas partitioning and a Big1-like H_2 SF recipe. Similar to our results, the gas phase metallicities in their models at a given luminosity are lower than the observational estimates taken at face value.

Fu et al. (2012) conducted a similar study, investigating four SF recipes and separately studying two gas partitioning recipes, BR and KMT, within the framework of the models developed by Fu et al. (2010) based on the MPA Millennium SAMs (Croton et al. 2006; Guo et al. 2011). Their 'Bigiel' SF recipe is similar to our Big1 recipe, but depends on the H_2 fraction, steepening below a critical value. Their 'Kennicutt' recipe is similar to our KS recipe. Their 'Genzel' recipe contains a linear scaling of $\Sigma_{\rm SFR}$ with $\Sigma_{\rm H_2}$, but also scales as the inverse galaxy dynamical time (which is redshift dependent). They also consider the KMT SF recipe. Again their results are quite consistent with ours. The stellar and H_2 mass functions are almost identical for all models, while the largest difference between models is in the H I content. This is the same conclusion that we reach based on PST14 and this study. All of their models underproduce massive galaxies at high redshift ($z \gtrsim 2$), as we found with similar SF prescriptions to the ones they adopted (i.e. our Big1 SF recipe, see Fig. 7). Fu et al. (2012) do not show their $m_{\rm star}$–SFR relation, but they show that the different SF recipes can lead to substantial deviation between model predictions for the cosmic SFR density at high redshift ($z \gtrsim 3$–4), as we also find. Interestingly, Fu et al. (2012) find that their gas phase MZR is *too shallow* compared with observations – the opposite problem to the one we encounter in our models. This is probably due to their adopted scaling for stellar driven winds. They assume that the mass outflow rate $\dot{m}_{\rm out} \propto \dot{m}_*$, i.e. a fixed mass loading factor, while we assume that the mass loading factor $\dot{m}_{\rm out}/\dot{m}_*$ scales approximately with inverse circular velocity squared. This dependence of MZR slope on wind scaling parameters is well known (Peeples & Shankar 2011). They also find that the *redshift evolution* of the MZR is quite sensitive to the SF recipe adopted; in particular, recipes in which SFE scales with galaxy dynamical time predict very weak or no evolution in the MZR, while their models without dynamical time scalings predict stronger evolution. This is qualitatively very similar to the weak evolution in the MZR seen in our fiducial GK and BR models (the non-linear slope of the Big2 recipe functions in a similar manner as scaling with dynamical time), and the much stronger evolution seen in our GK+Big1 models (see our Fig. 16).

Krumholz & Dekel (2012) implemented an updated version of the KMT recipe within simplified SAMs that only follow the mass accretion history of the main branch, and do not track the full merger trees. In their fiducial model they additionally assume a constant mass loading factor for stellar-driven winds (no dependence on galaxy mass or circular velocity) and strongly MEW. When we implement similar ingredients in our models (except that we retain our 'energy-driven' stellar wind scalings), we obtain qualitatively similar results. Namely, a metallicity-dependent formation efficiency for H_2 and self-consistent H_2-based SF recipe can significantly suppress and delay SF and metal enrichment in very low-mass haloes. We find that this effect is considerably stronger in our KMT+MEW models than in our fiducial GK models. This is because (a) in our fiducial GK model, the effect of a varying UV background partially compensates for the metallicity dependence of H_2-formation; and (b) the MEW delay enrichment of the cold gas, further suppressing SF (see Section 3.1). However, we stress that a noticeable effect is seen only in haloes with virial mass $M_{\rm h} \lesssim 10^{10.5}\,{\rm M}_{\odot}$ (stellar masses $m_* \lesssim 10^8 {\rm M}_{\odot}$). Our implementation of the KMT gas partitioning recipe and SF recipe, including MEW, does not significantly reduce the number density of galaxies in the stellar mass range $10^9{\rm M}_{\odot} \lesssim m_* \lesssim 10^{10}{\rm M}_{\odot}$, which is the mass scale of the low-mass galaxy problem discussed earlier; see Fig. B2. Therefore, although Krumholz & Dekel (2012) do not show their predicted stellar mass functions or stellar fractions, it is likely that their model also still suffers from the overprediction of low-mass galaxies ($m_{\rm star} \sim 10^{9-10}\,{\rm M}_{\odot}$) at intermediate redshifts ($0.5 \lesssim z \lesssim 4$) that we have discussed extensively above. Certainly they see the same qualitative problems with predicted sSFR being too low at intermediate redshifts that we have described here.

Galaxies hosted by haloes in the strongly affected mass range have typical stellar masses of $m_* \lesssim 10^8\,{\rm M}_{\odot}$, and it is probably not feasible to obtain complete samples of these objects at high redshift with existing facilities. However, this strong suppression of SF in low-mass haloes could have implications for reionization, future observations with the *James Webb Space Telescope*, and stellar archaeology in local dwarf galaxies. Moreover, Berry et al. (2014) showed that the strong suppression of H_2 formation and SF in low-mass haloes predicted by the KMT-like models would lead to a large population of 'barren' haloes that never experience significant SF, and so are filled with H I (see also Kuhlen et al. 2012). Berry et al. (2014) found that the existence of this population is in apparent tension with observations of H I absorption systems at high redshift.

We found that the inclusion of MEW produces a steeper MZR, leading to an even greater tension with observations. Krumholz & Dekel (2012) also found that their models produce a steeper gas phase MZR than observations, in spite of their adopted constant mass loading factor which generally leads to shallower predicted MZR. Interestingly, Krumholz & Dekel (2012) find significant evolution in the gas phase MZR from $z \sim 2$–0 in their models, in better apparent agreement with observations.

5 SUMMARY AND CONCLUSIONS

We have presented new models of galaxy formation, set in the framework of cosmological merger trees, which include physically motivated recipes for the partitioning of cold gas into an atomic, molecular, and ionized phase. These models have several advantages: first, we can make explicit predictions for the atomic and molecular gas properties of galaxies over cosmic time, which can be directly confronted with observations from current and upcoming facilities. The first of these predictions, for H I and H_2 gas observed in emission, were presented in PST14, and predictions for H I gas observed in absorption were presented in Berry et al. (2014). Popping et al. (2014b) extended these models to predict sub-mm line emission from several atomic and molecular species. These predictions may be directly compared with observations from current and upcoming facilities such as ALMA (Popping et al. in preparation), and also used to plan future observations with these facilities. Secondly, we can implement more physically motivated H_2-based recipes for SF within our models. In this paper, we focused on the predictions for the stellar mass content, SFR, and metallicity of galaxies in our new models, and compared these with available observational estimates from $z \sim 0$–6.

We summarize our main conclusions below.

(i) We performed a number of tests of the robustness of our models to resolution and various parameter values and ingredients. We find that our code is very well converged with respect to variation of the mass resolution of our merger trees. In addition, our results do not depend sensitively on the assumed values of the metallicity of pre-enriched gas or the molecular hydrogen floor (within a reasonable range of values).

(ii) Accounting for a reservoir of ionized gas (due to an internal and external photoionizing radiation field) in our models, which is not allowed to form molecular hydrogen or stars, does not significantly change our predictions.

(iii) We explored the effect of adopting different H_2-based SF recipes. All the recipes we considered gave similar results for low-mass galaxies. Models that adopted a 'two-slope' recipe (Big2), which has a linear dependence of SFRD on molecular gas surface density below a critical value, steepening to $\Sigma_{\rm SFR} \propto \Sigma_{\rm H2}{}^2$ above a critical value of $\Sigma_{\rm H2}$, produced more efficient SF and metal enrichment in massive galaxies at high redshift. Models that implement the Big2 recipe appear to be in better agreement with current observational estimates of the number density of massive galaxies at high redshift.

(iv) We explored the effect of different recipes for partitioning gas into an atomic and molecular component. The metallicity-based GK recipe and the pressure-based BR recipe gave surprisingly similar results, perhaps because of the strong correlation between disc mid-plane pressure ($\propto\Sigma_*$, to first order) and metallicity in our models. The KMT and GKFUV recipes, which do not include a dependence on the FUV radiation background as our fiducial GK recipe does, predicted less efficient formation of H_2, less SF and metal enrichment at early times, and later stellar mass assembly. These differences are only noticeable, however, in very low-mass haloes ($\log(M_{\rm h}/{\rm M}_{\odot}) \lesssim 10.5$).

(v) Both of our new fiducial models (GK and BR) reproduce the curvature in the relationship between total cold gas surface density and $\Sigma_{\rm SFR}$ seen in observations of nearby spiral galaxies. Our results illustrate the difficulty of disentangling the physical processes that are responsible for the scatter in this relationship, however, because of the strong correlations in galaxy properties.

(vi) The stellar mass function and mean stellar fractions ($f_{\rm star} \equiv m_*/M_{\rm h}$) of galaxies in our three fiducial models (the 'classic' KS model, the metallicity-based GK model and the pressure-based BR model) are almost identical for low-mass galaxies at all redshift $z \sim 0$–6. Both of the new models (GK and BR) predict earlier formation of massive galaxies, in better agreement with observational estimates than models that adopt the KS recipe. However, all the models investigated here still suffer from the overproduction of low-mass galaxies at intermediate redshift that has been noted in many previous works.

(vii) Although the median values of $f_{\rm star}$ are very similar in low-mass haloes in all three models, the models can have significantly differently shaped *distribution functions* $P(f_{\rm star}|M_{\rm h})$. In particular, the GK model tends to have a much more pronounced tail to low values of $f_{\rm star}$ in low-mass haloes ($\log(M_{\rm h}/{\rm M}_{\odot}) \lesssim 10.5$). The predicted broadening in $f_{\rm star}$ has potentially important implications for halo occupation models, which generally assume a narrow and fixed scatter in $f_{\rm star}(M_{\rm h})$.

(viii) All three fiducial models produce nearly identical predictions for the relationship between stellar mass and SFR at all redshifts. Even the distributions of sSFR at a given stellar mass are very similar. The KS model predicts a slightly narrower distribution of sSFR in high-mass galaxies at $z \gtrsim 1$ than the other two models. All the models investigated here predict SFR that are too low compared with observational estimates for low-mass galaxies at intermediate redshift.

(ix) All three models predict a weak dependence of the gas depletion time ($t_{\rm dep} \equiv (m_{\rm H\,I} + m_{\rm H_2})/\dot{m}_*$) on stellar mass, in agreement with observations of nearby normal disc galaxies and empirical estimates. The predicted $t_{\rm dep}$ decreases by about 1.3 dex from $z \sim 0$ to 6. The KS model predicts milder evolution in the depletion time for massive galaxies to high redshift, resulting in longer depletion times in massive high-redshift galaxies compared to the other two models.

(x) All three models predict quite good agreement with the observed $z = 0$ stellar mass versus metallicity relation (MZR) for *stellar* metallicities, but predict a gas phase MZR that is much steeper than observational estimates taken at face value. However, this tension may *perhaps* be relieved by properly accounting for the possible dependence of [α/Fe] on galaxy properties. The KS model predicts later metal enrichment of massive galaxies, leading to a shallower MZR at high redshift. In tension with observational results, all of our models predict a nearly constant or slightly declining metallicity for galaxies selected at fixed stellar mass from $z \sim 4$–0.

ACKNOWLEDGEMENTS

We thank Molly Peeples, Mark Krumholz, Fabian Walter, Adam Leroy, Nick Gnedin, and Joop Schaye for helpful discussions. We also thank Peter Behroozi and Paula Santini for providing their data in electronic form. In addition we thank the referee, Bruno Henriques, for comments and suggestions that improved the manuscript. RSS gratefully thanks the Downsbrough family for their generous support. This work has been supported in part by grant HST AR-13270.01 from NASA and a Simons Investigator award from the Simons Foundation.

REFERENCES

Andrews B. H., Martini P., 2013, ApJ, 765, 140
Arrigoni M., Trager S. C., Somerville R. S., Gibson B. K., 2010, MNRAS, 402, 173
Baldry I. K., Glazebrook K., Driver S. P., 2008, MNRAS, 388, 945
Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191
Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57
Berry M., Somerville R. S., Haas M. R., Gawiser E., Maller A., Popping G., Trager S. C., 2014, MNRAS, 441, 939
Bigiel F., Blitz L., 2012, ApJ, 756, 183
Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2846
Bigiel F. et al., 2011, ApJ, 730, L13
Blitz L., Rosolowsky E., 2004, ApJ, 612, L29
Blitz L., Rosolowsky E., 2006, ApJ, 650, 933 (BR)
Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517
Blumenthal G., Faber S. M., Flores R., Primack J. R., 1986, ApJ, 301, 27
Bondi H., 1952, MNRAS, 112, 195
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Boylan-Kolchin M., Ma C.-P., Quataert E., 2008, MNRAS, 383, 93
Brennan R. et al., 2015, MNRAS, 451, 2933
Bromm V., Larson R. B., 2004, ARA&A, 42, 79
Caffau E., Ludwig H.-G., Steffen M., Freytag B., Bonifacio P., 2011, Sol. Phys., 268, 255
Caputi K. I., Cirasuolo M., Dunlop J. S., McLure R. J., Farrah D., Almaini O., 2011, MNRAS, 413, 162
Chabrier G., 2003, PASP, 115, 763
Christensen C., Quinn T., Governato F., Stilp A., Shen S., Wadsley J., 2012, MNRAS, 425, 3058
Croton D. J. et al., 2006, MNRAS, 365, 11
Daddi E. et al., 2010, ApJ, 714, L118
Davé R., Finlator K., Oppenheimer B. D., 2012, MNRAS, 421, 98 (DFO12)
De Lucia G., Blaizot J., 2007, MNRAS, 375, 2

Dekel A., Mandelker N., 2014, MNRAS, 444, 2071
Dekel A., Zolotov A., Tweed D., Cacciato M., Ceverino D., Primack J. R., 2013, MNRAS, 435, 999
Duncan K. et al., 2014, MNRAS, 444, 2960
Dutton A. A., van den Bosch F. C., Dekel A., 2010, MNRAS, 405, 1690
Flores R., Primack J. R., Blumenthal G., Faber S. M., 1993, ApJ, 412, 443
Fontanot F., De Lucia G., Monaco P., Somerville R. S., Santini P., 2009, MNRAS, 397, 1776
Fu J., Guo Q., Kauffmann G., Krumholz M. R., 2010, MNRAS, 409, 515
Fu J., Kauffmann G., Li C., Guo Q., 2012, MNRAS, 424, 2701
Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, MNRAS, 362, 41
Genzel R. et al., 2010, MNRAS, 407, 2091
Glover S., 2013, in Wiklind T., Mobasher B., Bromm V., eds, Astrophysics and Space Science Library, Vol. 396, The First Galaxies. Springer-Verlag, Berlin, p. 103
Glover S. C. O., Clark P. C., 2014, MNRAS, 437, 9
Gnedin N. Y., 2000, ApJ, 542, 535
Gnedin N. Y., 2012, ApJ, 754, 113
Gnedin N. Y., Kravtsov A. V., 2010, ApJ, 714, 287
Gnedin N. Y., Kravtsov A. V., 2011, ApJ, 728, 88 (GK)
Greif T. H., Glover S. C. O., Bromm V., Klessen R. S., 2010, ApJ, 716, 510
Guo Q. et al., 2011, MNRAS, 413, 101
Haas M. R., Schaye J., Booth C. M., Dalla Vecchia C., Springel V., Theuns T., Wiersma R. P. C., 2013, MNRAS, 435, 2931
Haiman Z., Rees M. J., Loeb A., 1996, ApJ, 467, 522
Helfer T. T., Thornley M. D., Regan M. W., Wong T., Sheth K., Vogel S. N., Blitz L., Bock D., 2003, ApJS, 145, 259
Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R. E., Guo Q., Lemson G., Springel V., 2013, MNRAS, 431, 3373
Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, MNRAS, 451, 2663
Hirschmann M., Naab T., Somerville R. S., Burkert A., Oser L., 2012a, MNRAS, 419, 3200
Hirschmann M., Somerville R. S., Naab T., Burkert A., 2012b, MNRAS, 426, 237
Hopkins P. F. et al., 2009a, MNRAS, 397, 802
Hopkins P. F., Cox T. J., Younger J. D., Hernquist L., 2009b, ApJ, 691, 1168
Katz N., 1992, ApJ, 391, 502
Kennicutt R. C., Jr, 1989, ApJ, 344, 685
Kennicutt R. C., Jr, 1998, ApJ, 498, 541
Kewley L. J., Ellison S. L., 2008, ApJ, 681, 1183
Khaire V., Srianand R., 2015, ApJ, 805, 33
Kimm T. et al., 2009, MNRAS, 394, 1131
Kirby E. N., Lanfranchi G. A., Simon J. D., Cohen J. G., Guhathakurta P., 2011, ApJ, 727, 78
Kirby E. N., Cohen J. G., Guhathakurta P., Cheng L., Bullock J. S., Gallazzi A., 2013, ApJ, 779, 102
Komatsu E. et al., 2009, ApJS, 180, 330
Komatsu E. et al., 2011, ApJS, 192, 18
Kravtsov A. V., 2013, ApJ, 764, L31
Kravtsov A. V., Klypin A. A., Khokhlov A. M., 1997, ApJS, 111, 73
Kravtsov A. V., Gnedin O. Y., Klypin A. A., 2004, ApJ, 609, 482
Krumholz M. R., Dekel A., 2012, ApJ, 753, 16
Krumholz M. R., Gnedin N. Y., 2011, ApJ, 729, 36
Krumholz M. R., McKee C. F., Tumlinson J., 2008, ApJ, 689, 865
Krumholz M. R., McKee C. F., Tumlinson J., 2009a, ApJ, 693, 216
Krumholz M. R., McKee C. F., Tumlinson J., 2009b, ApJ, 699, 850
Krumholz M. R., Dekel A., McKee C. F., 2012, ApJ, 745, 69
Kuhlen M., Krumholz M. R., Madau P., Smith B. D., Wise J., 2012, ApJ, 749, 36
Lagos C. D. P., Baugh C. M., Lacey C. G., Benson A. J., Kim H.-S., Power C., 2011a, MNRAS, 418, 1649
Lagos C. D. P., Lacey C. G., Baugh C. M., Bower R. G., Benson A. J., 2011b, MNRAS, 416, 1566 (L11a)
Lagos C. d. P., Bayet E., Baugh C. M., Lacey C. G., Bell T. A., Fanidakis N., Geach J. E., 2012, MNRAS, 426, 2142
Lang P. et al., 2014, ApJ, 788, 11
Leroy A. K., Walter F., Brinks E., Bigiel F., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2782
Leroy A. K. et al., 2009, AJ, 137, 4670
Leroy A. K. et al., 2013, AJ, 146, 19
Lu Y. et al., 2014, ApJ, 795, 123
Madau P., Dickinson M., 2014, ARA&A, 52, 415
Martin C. L., Kennicutt R. C., 2001, ApJ, 555, 301
Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319
Moustakas J. et al., 2013, ApJ, 767, 50
Muñoz J. A., Peeples M. S., 2015, MNRAS, 448, 1430
Murray N., Rahman M., 2010, ApJ, 709, 424
Muzzin A. et al., 2013, ApJ, 777, 18
Nakamura F., Umemura M., 2001, ApJ, 548, 19
Narayanan D., Krumholz M. R., Ostriker E. C., Hernquist L., 2012, MNRAS, 421, 3127
Niemi S.-M., Somerville R. S., Ferguson H. C., Huang K.-H., Lotz J., Koekemoer A. M., 2012, MNRAS, 421, 1539
Obreschkow D., Rawlings S., 2009, MNRAS, 394, 1857
Ostriker E. C., McKee C. F., Leroy A. K., 2010, ApJ, 721, 975
Parkinson H., Cole S., Helly J., 2008, MNRAS, 383, 557
Peeples M. S., Shankar F., 2011, MNRAS, 417, 2962
Peeples M. S., Werk J. K., Tumlinson J., Oppenheimer B. D., Prochaska J. X., Katz N., Weinberg D. H., 2014, ApJ, 786, 54
Planck Collaboration XIII 2015, preprint (arXiv:1502.01589)
Popping G., Somerville R. S., Trager S. C., 2014a, MNRAS, 442, 2398 (PST14)
Popping G., Pérez-Beaupuits J. P., Spaans M., Trager S. C., Somerville R. S., 2014b, MNRAS, 444, 1301
Popping G., Behroozi P. S., Peeples M. S., 2015, MNRAS, 449, 477
Porter L. A., Somerville R. S., Primack J. R., Johansson P. H., 2014, MNRAS, 444, 942
Robertson B. E., Kravtsov A. V., 2008, ApJ, 680, 1083
Robertson B., Cox T. J., Hernquist L., Franx M., Hopkins P. F., Martini P., Springel V., 2006, ApJ, 641, 21
Robitaille T. P., Whitney B. A., 2010, ApJ, 710, L11
Saintonge A. et al., 2011, MNRAS, 415, 32
Salim S. et al., 2007, ApJS, 173, 267
Salmon B. et al., 2015, ApJ, 799, 183
Santini P. et al., 2012, A&A, 538, A33
Schaye J., 2004, ApJ, 609, 667
Schaye J. et al., 2010, MNRAS, 402, 1536
Schmidt M., 1959, ApJ, 129, 243
Schmidt M., 1963, ApJ, 137, 758
Schneider R., Ferrara A., Natarajan P., Omukai K., 2002, ApJ, 571, 30
Schörck T. et al., 2009, A&A, 507, 817
Schruba A. et al., 2011, AJ, 142, 37
Sharon C. E., Baker A. J., Harris A. I., Thomson A. P., 2013, ApJ, 765, 6
Somerville R. S., Davé R., 2015, ARA&A, 53, 51
Somerville R. S., Kolatt T. S., 1999, MNRAS, 305, 1
Somerville R. S., Primack J. R., 1999, MNRAS, 310, 1087
Somerville R. S., Primack J. R., Faber S. M., 2001, MNRAS, 320, 504
Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008a, MNRAS, 391, 481 (S08)
Somerville R. S. et al., 2008b, ApJ, 672, 776
Somerville R. S., Gilmore R. C., Primack J. R., Domínguez A., 2012, MNRAS, 423, 1992 (S12)
Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, ApJS, 214, 15
Steinhardt C. L. et al., 2014, ApJ, 791, L25
Stoll R., Prieto J. L., Stanek K. Z., Pogge R. W., 2013, ApJ, 773, 12
Sutherland R., Dopita M. A., 1993, ApJS, 88, 253
Thomas D., Maraston C., Bender R., Mendes de Oliveira C., 2005, ApJ, 621, 673
Tolstoy E., Hill V., Tosi M., 2009, ARA&A, 47, 371
Tomczak A. R. et al., 2014, ApJ, 783, 85
Toomre A., 1964, ApJ, 139, 1217
Trager S. C., Faber S. M., Worthey G., González J. J., 2000, AJ, 119, 1645

Walter F., Brinks E., de Blok W. J. G., Bigiel F., Kennicutt R. C., Thornley M. D., Leroy A., 2008, AJ, 136, 2563
Weinmann S. M., Pasquali A., Oppenheimer B. D., Finlator K., Mendel J. T., Crain R. A., Macciò A. V., 2012, MNRAS, 426, 2797
Whitaker K. E. et al., 2014, ApJ, 795, 104
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
White C. E., Somerville R. S., Ferguson H. C., 2015, ApJ, 799, 201
Wise J. H., Turk M. J., Norman M. L., Abel T., 2012, ApJ, 745, 50
Wong T., Blitz L., 2002, ApJ, 569, 157
Woo J., Courteau S., Dekel A., 2008, MNRAS, 390, 1453
Yates R. M., Henriques B., Thomas P. A., Kauffmann G., Johansson J., White S. D. M., 2013, MNRAS, 435, 3500
Zahid H. J., Geller M. J., Kewley L. J., Hwang H. S., Fabricant D. G., Kurtz M. J., 2013, ApJ, 771, L19

APPENDIX A: ADDITIONAL TESTS

In Fig. A1, we test the impact of varying the mass resolution of our merger trees. Note that what we mean by the 'mass resolution' here is the mass of the smallest haloes that are tracked in the merger tree. This is not equivalent to the particle mass in an *N*-body simulation, but rather to the smallest halo mass that can be robustly identified. We see from this test that in order to robustly reconstruct the whole halo mass accretion history back to $z \sim 10$, we require a minimum halo mass of $\sim 1/100$ of the root mass at the output redshift. Accordingly, we impose this condition on all haloes in the runs used in this work. It is reassuring to see that, once the halo mass accretion history is well resolved, our SAM predictions converge extremely

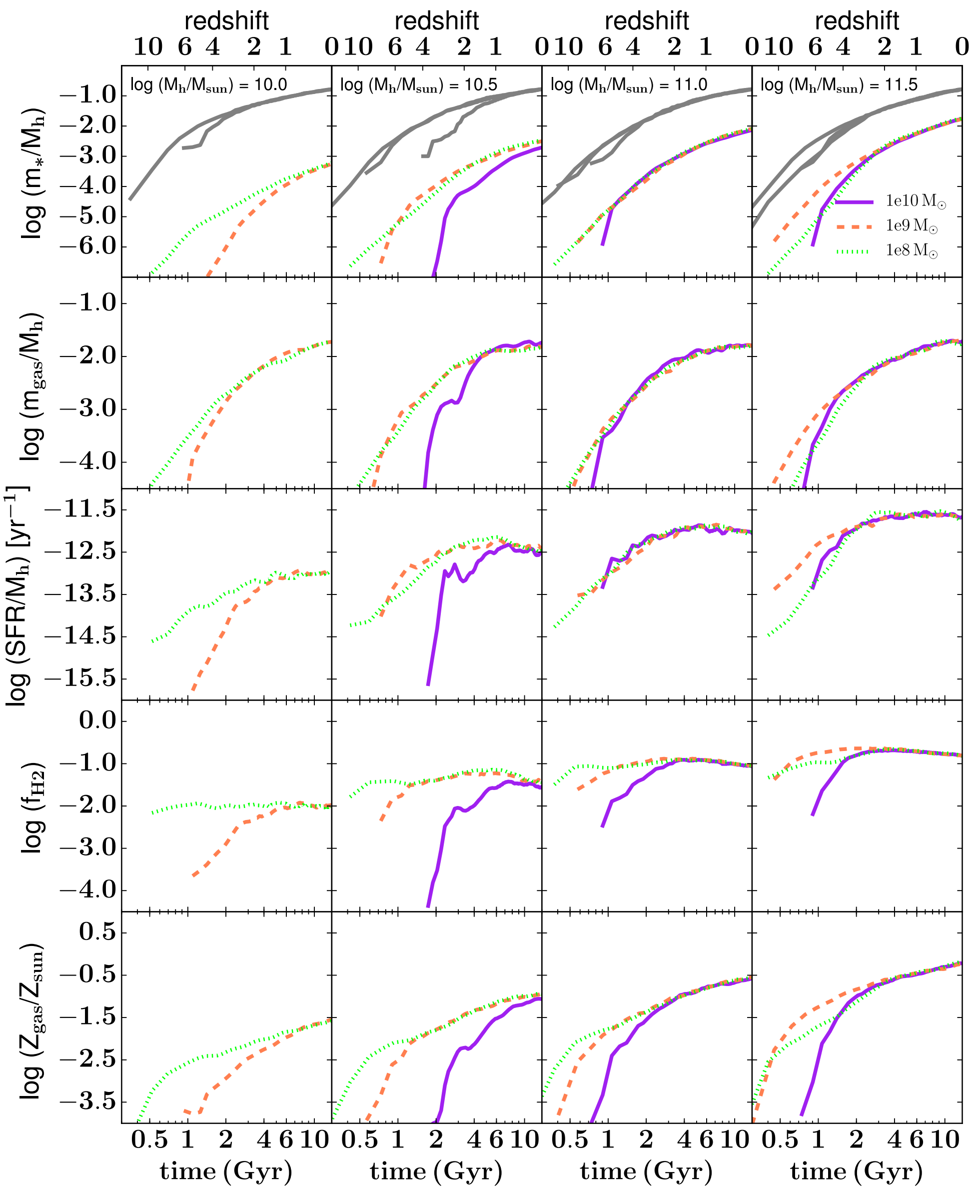

Figure A1. From top to bottom, coloured lines show the stellar mass, cold neutral gas mass (H I +H_2), and SFR normalized by the mass of the root halo at $z = 0$ for the largest progenitor galaxy as a function of cosmic time (redshift). The H_2 fraction ($f_{H_2} \equiv m_{H_2}/(m_{H_2} + m_{H\,I})$) and gas phase metallicity (in solar units) are also shown. Grey lines show the maximum baryon fraction in the halo, $f_b\, M_h(t)$, where $M_h(t)$ is the mass of the largest progenitor halo at time t and f_b is the universal baryon fraction (different grey lines correspond to different resolutions; lower resolution runs cannot resolve the halo mass accretion history as far back in time). In this experiment, we test the dependence of our results on the mass resolution of our merger trees, varying the mass resolution by two orders of magnitude. Results are shown for a mass resolution of 10^{10} M$_\odot$ (solid purple), 10^9 M$_\odot$ (dashed orange), and 10^8 M$_\odot$ (dotted green). The mass accretion histories are well resolved when the mass resolution is at least 0.01 times the mass of the root halo.

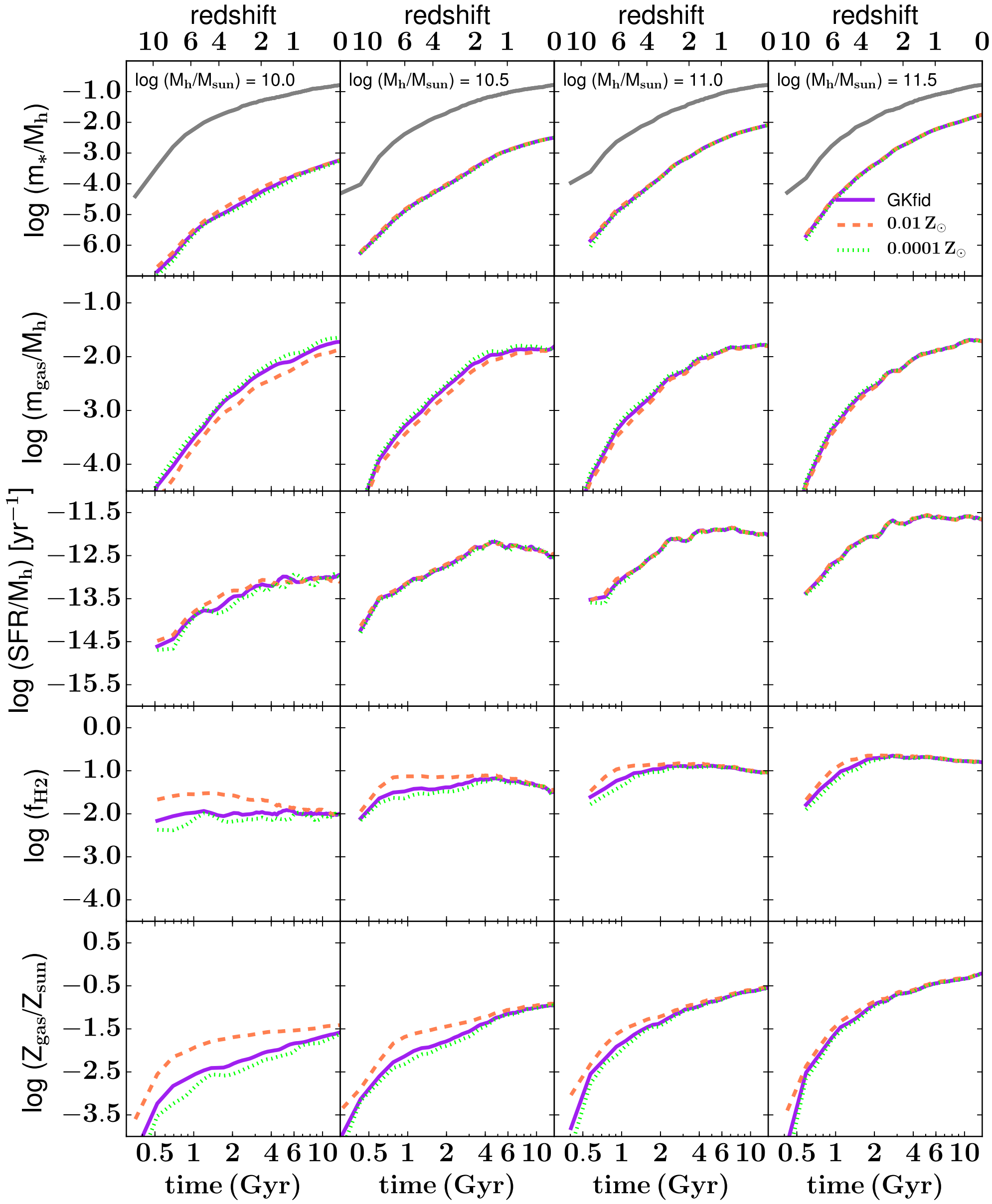

Figure A2. Same as Fig. A1, except here we compare the results of our fiducial GK model with different values for the 'pre-enriched' gas metallicity ($Z_{\mathrm{pre-enrich}}$), as shown in the key. Our results are quite insensitive to the value of this parameter within a reasonable range.

well (note that we do not retune the free parameters when we change the mass resolution).

Next we test for possible sensitivity to the values of two parameters that we introduce to simulate the formation of stars in primordial gas, the metallicity of the 'pre-enriched' gas $Z_{\mathrm{pre-enrich}}$ (most relevant for the metallicity-based recipes), and the primordial molecular hydrogen fraction $f_{\mathrm{H_2,floor}}$ (most relevant for the BR recipe). Leaving all other settings of the fiducial GK model fixed, we vary $Z_{\mathrm{pre-enrich}}$ from its fiducial value of 10^{-3} by one order of magnitude downwards, to 10^{-4}, and upwards to 0.01 (see Fig. A2). Overall, the impact of even such extreme variations is minor. The most noticeable impact is on the stellar and gas phase metallicities. The metallicity builds up earlier in models with higher values of $Z_{\mathrm{pre-enrich}}$, as expected. As a result, the H_2 fraction is higher at earlier times in the model with higher $Z_{\mathrm{pre-enrich}}$, leading to higher SFE and slightly lower gas fractions. Similarly, we varied the value of $f_{\mathrm{H_2,floor}}$ from its fiducial value of 10^{-4} up and down by an order of magnitude. This has no discernable effect on our results, and we therefore omit the corresponding figure.

In a related experiment, we run our (otherwise) fiducial GK model with metal-enriched winds, described in Section 2.4. Here, the metallicity of stellar-driven outflows can be metal enhanced relative to the ISM by a factor that depends on the halo mass. As seen in Fig. A3, we find that MEW can significantly delay the formation of H_2 and stars in very low mass haloes ($\log M_{\mathrm{h}}/\mathrm{M}_{\odot} = 10.0$), and cause the build-up of slightly more cold gas in low-mass haloes ($\log M_{\mathrm{h}}/\mathrm{M}_{\odot} \lesssim 10.5$); note however that these haloes host galaxies that are well below the detection limits of most surveys except in the very nearby Universe ($m_* \simeq 10^7$–$10^8\,\mathrm{M}_{\odot}$). Metal-enriched winds can also delay the build-up of metal-enriched gas even in more massive haloes ($\log M_{\mathrm{h}}/\mathrm{M}_{\odot} \lesssim 11.5$).

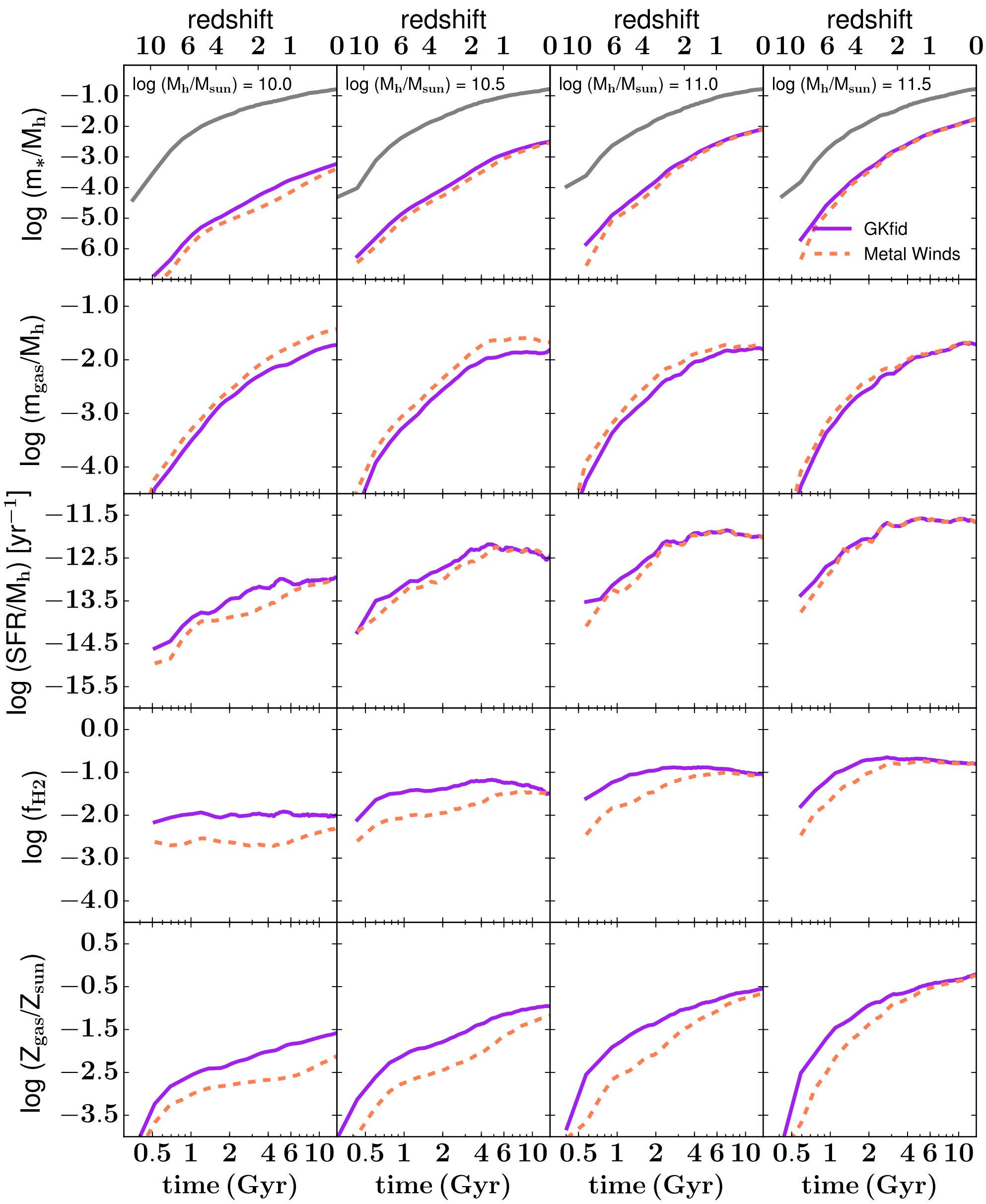

Figure A3. Same as Fig. A1, except here we compare the results of our fiducial GK model with and without MEW. MEW can delay enrichment and, to a lesser extent, H_2 and SF in low-mass haloes.

A new ingredient we have introduced into our models is the tracking of gas that is photoionized either by an external radiation field or by internal sources. This gas is not eligible to form H_2 or stars. In Fig. A4, we show the galaxy properties in the fiducial GK model with and without tracking of H II. Although our model predicts that galaxies contain a substantial amount of H II (see fig. 2 in PST14), partitioning this gas into a separate reservoir has a very weak effect on our results. The only noticeable effect is slightly lower H_2 fractions at high redshift, particularly in the lowest mass haloes.

APPENDIX B: RESULTS FOR ADDITIONAL MODEL VARIANTS

In this appendix we show supplementary results for several additional model variants, in order to aid the interpretation of the results presented in the main text. In Fig. B1, we show the stellar mass function at $z = 0$, 1, 2, and 6 for our fiducial GK model (the same one shown in Fig. 7), compared with the GK model with a fixed value of the UV radiation background $U_{\rm MW} = 1$ (GKFUV), the GK model for gas partitioning with the Big1 SF relation (GK+Big1), and a model with the KMT recipes for gas partitioning and a KMT SF relation (see Table 2). Fig. B2 shows the median stellar fraction as a function of halo mass for central galaxies, in the same suite of models, with the addition of the GK model with photoionization 'squelching' switched off shown in the $z = 0$ panel only.

These plots illustrate several points, which we already discussed in Section 3.1. First, the metallicity and UV radiation field dependence in the GK recipe partially counteract each other (lower mass galaxies have lower metallicity, resulting in less efficient H_2 formation, but also a lower SFR, resulting in less efficient H_2 destruction). Recipes that do not account for the effect of a varying UV radiation field (GKFUV and KMT) predict that H_2 formation becomes so

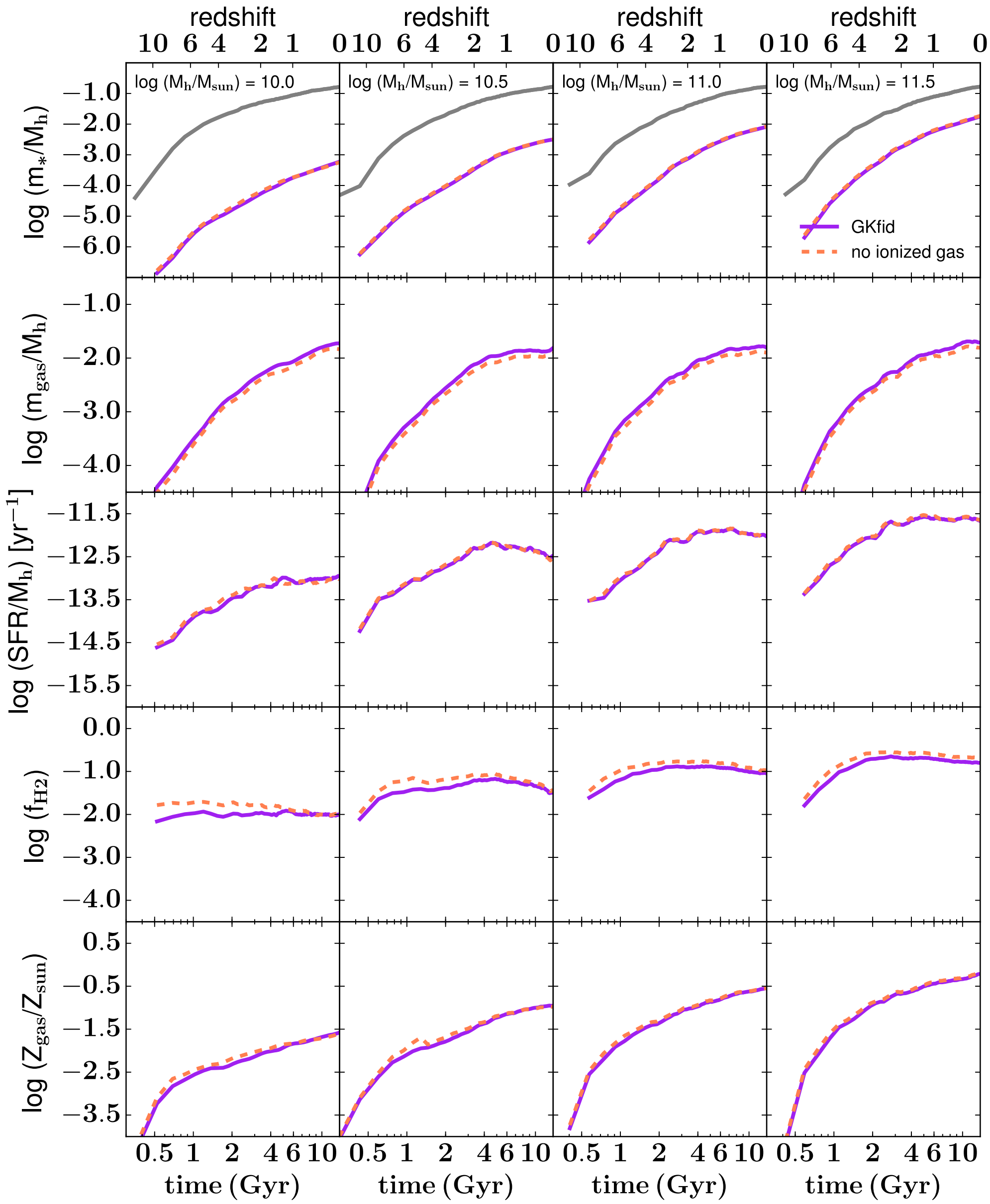

Figure A4. Same as Fig. A1, except here we compare the GK model with and without the partitioning of ionized gas (H II) into a separate reservoir. Our results are nearly unchanged whether or not we include the ionized gas component.

inefficient in low-mass galaxies that the stellar mass function actually turns over at $\log(m_*/\mathrm{M}_\odot) \simeq 8$. Similarly, $f_{\rm star}(M_{\rm h})$ declines sharply at $\log(M_{\rm h}/\mathrm{M}_\odot) \simeq 10$. Note that although in the models shown, the abundance of very low mass galaxies is actually probably too small compared with observations, we could probably adjust our recipes for stellar feedback or photoionization squelching to fix this. However, it does appear that even in these models, the excess of low-mass galaxies ($m_{\rm star} \sim 10^{9-10}\,\mathrm{M}_\odot$) at intermediate redshift ($0.5 \lesssim z \lesssim 4$) persists. This indicates that the halo mass scale where SF can become inefficient enough to break the self-regulation equilibrium is smaller than the one where the discrepancy with current observations appears.

The second point is that the differences between models seen in more massive haloes are almost solely due to the assumed scaling of the SF relation at large gas surface densities. The model with the steepest dependence of $\Sigma_{\rm SFR}$ on $\Sigma_{\rm H2}$ (Big2, with $N_{\rm SF} \rightarrow 2$ in dense gas) has the highest number density of massive galaxies and the highest values of $f_{\rm star}$ in massive haloes. The GKFUV model is almost identical to the fiducial GK model at high masses. The KMT model is the next highest ($N_{\rm SF} \rightarrow 1.4$ in dense gas), then the GK+Big1 model ($N_{\rm SF} = 1$). This is because regardless of the gas partitioning recipe, gas in these galaxies is dense enough that it is nearly all molecular. Stellar-driven winds cannot efficiently escape these deep potential wells. Therefore there is a strong dependence on the gas depletion time (SFE).

It is also clear from Fig. B2 that modelling of photoionization squelching will have an extremely degenerate effect with that of gas partitioning and stellar feedback on stellar properties. However, observations of gas content should help break these degeneracies.

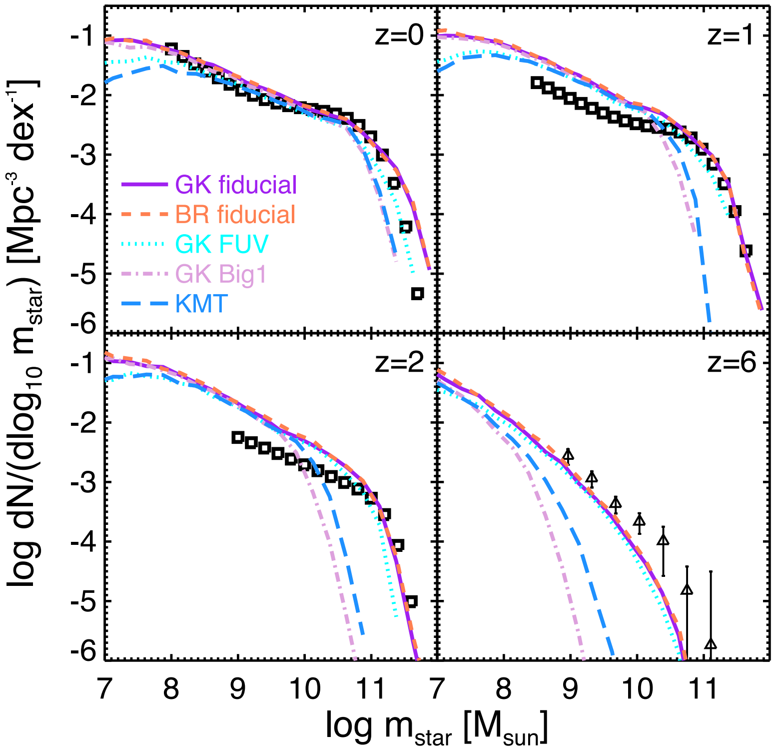

Figure B1. Stellar mass function evolution with redshift. Symbols show observational estimates as detailed in Fig. 7. The purple solid line shows the results of the fiducial GK model, the cyan dotted line shows the GKFUV model, the dashed lavender line shows the GK+Big1 model and the long-dashed blue line shows the KMT model (see text and Table 2).

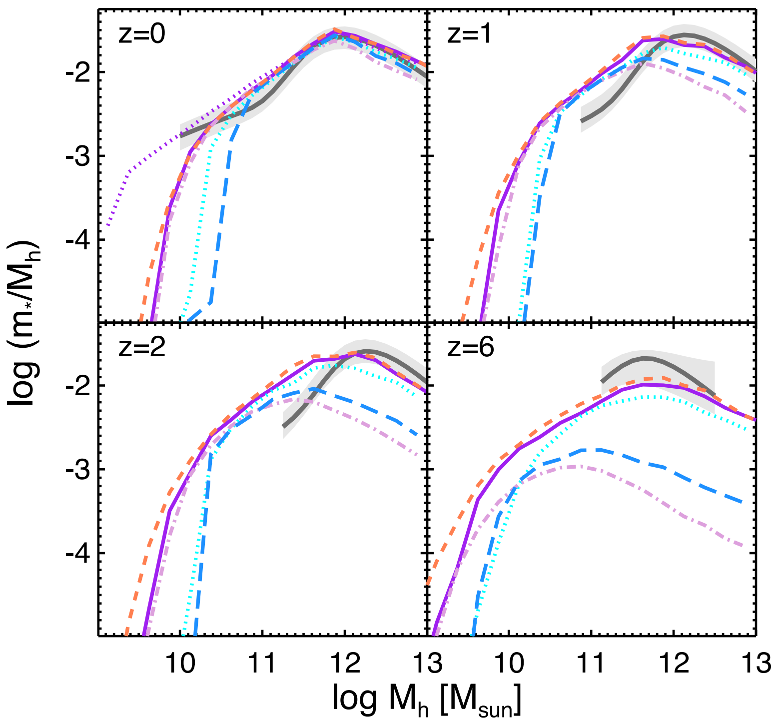

Figure B2. The stellar mass of central galaxies divided by the total mass of their dark matter halo, in redshift bins from $z = 0$–6. Dark grey solid lines and shaded areas show constraints from abundance matching from Behroozi et al. (2013). Models shown are as in Fig. B1. In addition, the dot–dashed purple line (shown in the $z = 0$ panel only) shows the fiducial GK model with photoionization squelching switched off.

This paper has been typeset from a TEX/LATEX file prepared by the author.